# Design and Implementation of Sound Projector Using a 30-Element Loudspeaker

Jordi Alvaro Arqués

Examiner: Prof. Per Zetterberg

November 19, 2009


# Abstract

For many years, the surround sound effect could only be fully experienced by setting up several speakers, which required considerable wiring for installation in a typical room. Digital sound projectors, one of the most sought-after state-of-the-art solutions, use an array of small speakers enclosed in a box to provide surround by creating virtual sound sources around the listener.

This thesis presents a DSP implementation of a simple digital sound projector based on a uniform linear array of thirty loudspeakers. Surround is produced by controlling separate beams of sound. These are reflected off the walls and arrive at the listener's position from the desired directions. Due to hardware constrains, the bandwidth of the sound is limited.

In order to estimate the direction of the sound at the listener, two techniques based on estimating the time difference of arrival (TDOA), Classical Cross-Correlation (CC) and Cross-Correlation with the Source Signal (CSS), are compared. Objective test results show that the CSS technique is more robust than CC under noisy environments, while both techniques behave similarly in quiet conditions.

Subjective and objective results show that this implementation of a sound projector is successful, even if front virtual channels are better perceived by the listener than the rear channels.

# Acknowledgments

This master's thesis was carried out at the School of Electrical Engineering, Kungliga Tekniska Högskolan in Stockholm, Sweden. It is the final part of my education at KTH and the Universidad Politècnica de Barcelona, Spain.

I would like to express my gratitude to my examinor Per Zetterberg for the opportunity of working on this thesis and for the unconditional guidance and assistance he provided to me during its development.

I would also like to thank to my friend, Antonio Campos, for our helpful conversations and advice.

Finally, I would like to thank my beloved parents for the never-ending support, help and invaluable time they invested in me. Without their love and support I would never have succeeded at these studies.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| **A/D** | Analog-to-digital converter |
| **AF** | Array factor |
| **BP** | Band-pass |
| **C** | C programming language |
| **CC** | Cross-correlation method |
| **CCS** | Code Composer Studio |
| **CSS** | Cross-correlation with the source signal |
| **GCC** | Generalized cross-correlation method |
| **D/A** | Digital-to-analog converter |
| **DMA** | Direct Memory Access |
| **DOA** | Direction of arrival |
| **DSP** | Digital signal processor |
| **DVD** | Digital Versatile Disk |
| **FIR** | Finite impulse response |
| **GUI** | Graphical user interface |
| **I/O** | Input/output |
| **PC** | Personal computer |
| **TDOA** | Time difference of arrival |
| **ULA** | Uniform linear array |
| **USB** | Universal Serial Bus |

# Chapter 1

# Introduction

## 1.1 Background and Problem Statement

Home theaters bring us the opportunity to enjoy the experience of watching and listening to movies, sports, concerts, music and even games in an improved way. Modern technology allows us to create a specific and good environment, but it needs the use of several components.

Focusing on sound perception, in order to create truly dramatic soundscapes, in general some speakers and a subwoofer placed around the room are required. In fact, different techniques, such as 5.1-channel surround (five speakers and one subwoofer) or 7.1-channel surround (seven speakers and one subwoofer), are used.

**Figure 1.1.** 5.0-channel setup using a loudspeaker array.

The positioning of the speakers is one of the most important factors to get surround sound right. A typical 5.1-channel setup involves a pair of front speakers, a centre channel speaker, a pair of rear or surround speakers and a subwoofer (a loudspeaker specialized in low frequencies).

Premium surround sound systems are a dream for almost everyone, but many are reluctant to implement a real 5-channel sound system at home. The reason is the hassle of placing all five speakers in one room and connecting them with wires.

Digital sound projectors are one of the most sought-after state-of-the-art solutions. They provide true multi-channel surround sound from only one component. The projector utilizes an array of many small loudspeakers to form a number of virtual speakers and mimic multichannel sound. The secret of these virtual speakers is multiple beamforming, so different angles-of-arrival, powers and delays are achieved. By producing tight and focusable beams of sound, the digital sound projector transmits the separate sound channels around the listener's room. By reflecting off walls and other surfaces, these beams finally arrive at the listener from the chosen directions. Figure 1.1 shows a surround sound system using a loudspeaker array.

Nowadays, some important and well-recognized brands (Yamaha, Pioneer, Loewe…) have developed commercial sound projectors which can produce a surround effect in a specific room.

The quality and prices vary a lot among them: while some are very expensive and deliver a good-quality 5.1-channel setup, other cheaper systems can only produce a simple sensation of being surrounded by the sound but in which it is not possible to feel it coming from a specific direction [1].

**Figure 1.2.** Example of current commercial sound projectors. Models are Yamaha's YSP-4000 (left) and Pioneer's PDSP-1 (right).

## 1.2 Literature Related to Array Beamforming

Several techniques have been developed in the field of array beamforming. In the following, an overview of the methods focused on steering the main lobe towards a specific direction is given [2][3].

- **Phased array and time-delay techniques.** They consist on focusing the main lobe towards a specific direction by delaying samples or the phase of the signals in each transducer.
- **Fourier and Woodyard-Lawson syntheses.** They are beam-shaping methods in which the radiation pattern is specified.
- **Chebychev and Taylor syntheses.** They are used in radiation patterns with narrow main lobes and low powered sidelobes.
- **Adaptive beamformers** [4]**.** The Minimum Variance Distortionless Response (MVDR) and the Generalized Sidelobe Canceller (GSLC) [5] are included here.

Even though it is known that the phased array is an optimal beamformer only when there is no interference, it is the most popular algorithm of those named above.

## 1.3 Literature Related to DOA and TDOA Estimation

The direction of arrival (DOA) and beamforming algorithms in the sound field have been developed and studied in depth during the last decades [6].

Radars and sonar are the main applications of these methods. However, there are also some other applications in the sound field; for example, in video-conferencing systems, the fact of knowing the direction from where a speaker is talking permits to incorporate video-cameras which can be steered automatically towards that person; moreover, in the development of a humanoid robot, speech signal processing becomes a technology of much interest to the research field of human-robot interaction [7].

The main techniques used in sound applications are based on obtaining the time difference of arrival (TDOA) [8][9][10] and are summed up in the following subsections.

### 1.3.1 The Family of the Generalized Cross-Correlation Methods (GCC)

This algorithm was proposed by Knapp and Carter [11]. It is based on obtaining the inverse discrete Fourier transform (IDFT) from the cross-spectrum of the sensors filtered by a frequency-domain weighting function. There are many options for the weighting function, leading to a variety of methods.

- **Classical Cross-Correlation.** It obtains the TDOA from the usual cross-correlation function. In this case the weighting function is equal to one.
- **Smoothed Coherence Transform (GCC-SCOT)** [12]**.** This case is used to overcome the impact of fluctuating levels of the source speech signal. It pre-whitens the sensor outputs before their cross-spectrum is calculated.

- **Phase Transform (GCC-PHAT)** [13][14][15]**.** This algorithm combats the room reverberation effects due to multi-path very well. It pre-whitens the signals using the information included in its phase.

### 1.3.2 High-Resolution Spectral-Estimation-Based Locators

These locators include the modern techniques adapted from the field of high-resolution spectral analysis [3][4][6][16]: autoregressive (AR) modeling, minimum variance (MV) spectral estimation, and the variety of eigenanalysis-based techniques (of which the MUSIC algorithm is the most popular).

All these algorithms were originally developed for narrowband signals, but, recently, they have been extended to processing wideband signals [3][4][6]. These techniques are less robust to source and sensor modeling errors than the family of GCC methods. Moreover, signal coherence, such as that created by reverberating conditions, is detrimental to algorithmic performance [9].

## 1.4 Model of the System

The goal of this project is to design and develop the software for a DSP in order to perform a digital sound projector using a 30-element loudspeaker array, 4 omnidirectional microphones, a DSP card and a DVD-player. The selected configuration is a broadband (0-20000 *Hz*) 4.0-channel setup (front right and left, and rear right and left channels). The sound has to surround the listener as if it arrives from the four directions indicated in Figure 1.3 below.

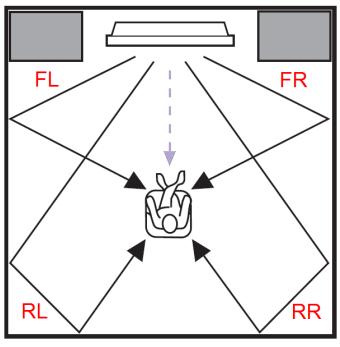



**Figure 1.3.** 4.0-channel setup using a loudspeaker array: front left (FL) and right (FR), and rear left (RL) and right (RR). The purple line shows the path followed by the direct sound which must be minimized or if possible removed.

Due to imperfections in the procedure of beamforming, some sidelobes are always received from the loudspeaker array. The sidelobes which affect the most are the ones directed towards the location of the listener (purple path in Figure 1.3). These sidelobes act as a second source of sound while the reflected sound beams, the ones steered towards the direction of interest, act as a primary source. The directional characteristic of the same sound arriving from two different sources depends on intensity difference and time delay between the two sound waves. An important issue lies in minimizing the effect produced by this direct sound.

The software of the system is composed by three parts:

- **Sounding phase:** Since the dimensions and the shape of the room are unknown, first it must be studied as follows. The sound projector transmits a signal towards a specific number of directions in the plane and a constellation of sensors, set at the position of the listener, store the received signals. The DOA of the signal received from the sensors is estimated. The procedure is repeated for several transmit directions. The solution adopted boils down to making a "transmit-receiver direction" table using both direction of transmission and arrival and the recorded sound power level. From now, we will refer to this table as sounding table. Then, the parameters are selected for each of the four channels.

- **Graphics phase:** This phase is optional and is not required in order to run the projector phase. Its purpose is to graphically show the results obtained from the sounding phase.

- **Projector phase:** Once the transmit parameters have been obtained, the system is ready to play the music or sound received from the DVD-player through the four channels in real-time. This is done by beamforming the signals aimed for different channels (i.e. front right and left, and rear right and left channels) and superimposing them at the input of the array. The listener must be sat at the position where the microphones were located during the sounding phase.

The Classical Cross-Correlation technique included in the GCC methods is the selected technique for DOA estimation. This choice was determined by knowing that the system has to work under reverberation and multiple-path effects: the fact that the aim of the sound projector is to play sound (music, the sound of a movie…) in a room provokes that some reflections of the sound waves appear from the right, left and rear walls. Therefore, those effects must be shown on the results in order to obtain a real DOA effect at the position of the listener.

The chosen method to steer the beam was the time-delay beamforming due to its simplicity. The real-time requirement was key for this decision, because of a fast algorithm was needed.

## 1.5 Outline

After this introductory chapter, the report is structured as follows.

Chapter 2, Description of the Work Elements, presents the selected system and the place where the simulations have been carried out. Next, Chapter 3, Methodology, gives the necessary

theoretical details to implement the methods on the DSP and MATLAB. Chapter 4, Signal Description, gives the details and features of the transmitted signal. Chapter 5, Implementation, explains the functionality of the programs and their structure. In Chapter 6, Results, an account of the experimental results of the thesis as well as their analysis and discussion are presented in several figures. Subjective listening tests and objective measures are registered from different points of view. Finally, Chapter 6, Conclusions and Future Work, sums up the main findings of this thesis and suggests several ways to extend it. Appendix A shows the possible directions of transmission at the beamformer. Appendix B contains the user manual to manage the GUI in MATLAB. Appendix C and Appendix D add supplementary tables and graphics obtained from the objective results.

# Chapter 2

# Description of the Work Elements

Features of both hardware and software used in the implementation are explained in this section in detail. Also, the room where the tests were performed and the coordinate system are described.

## 2.1 Test Room

The place chosen was a laboratory room located in the Electrical Engineering department. The chamber was furnished with typical laboratory material and its dimensions were 11 *m* x 11 *m*. A diagram of the floor plan layout and the positions of the microphones and the sound source are shown in Figure 2.1.

The sound projector was placed at a distance of 2 *m* from the East wall and 9 *m* from the Right wall. The distance between the loudspeaker array and the matrix of microphones was set up to 4.3 *m*, and the distance between the sensors and the South wall was 1.7 *m*.

## 2.2 Coordinate System

From now, all the comments about the laboratory room will refer to the floor plan shown in Figure 2.1. For convenience, the reference system is defined from the position where the listener must be located during the projector phase (reception point). This position coincides with the one of the microphones in the sounding phase.

The sound projector and the sensors need the use of different coordinate systems. Therefore, two coordinate patterns, one for the transmitter and the other for the receiver, must be introduced:

- The transmitter coordinate system only covers the half plane to which the speakers are focused. It is defined as Polar coordinate ranged from 0º to 180º. This angle is designated as $\varphi$. The zero degree is oriented towards the left wall (from the point of view of the listener), and the 90 degree towards the sensors.
- The receiver coordinate system covers the whole plane. It is defined in Polar and Cartesian coordinates.
    - The Polar coordinate is ranged from 0º to 360º. This angle is designated as $\theta$. The zero degree is oriented towards the center of the sound projector, and the 90 degree towards the left wall (from the point of view of the listener).
    - The Cartesian plane is defined from the center of the matrix of microphones designating the positive $x$-axis towards the 0º of the Polar coordinate and the positive $y$-axis towards the 90º.

These patterns and their orientation in the laboratory room are shown in Figure 2.1.

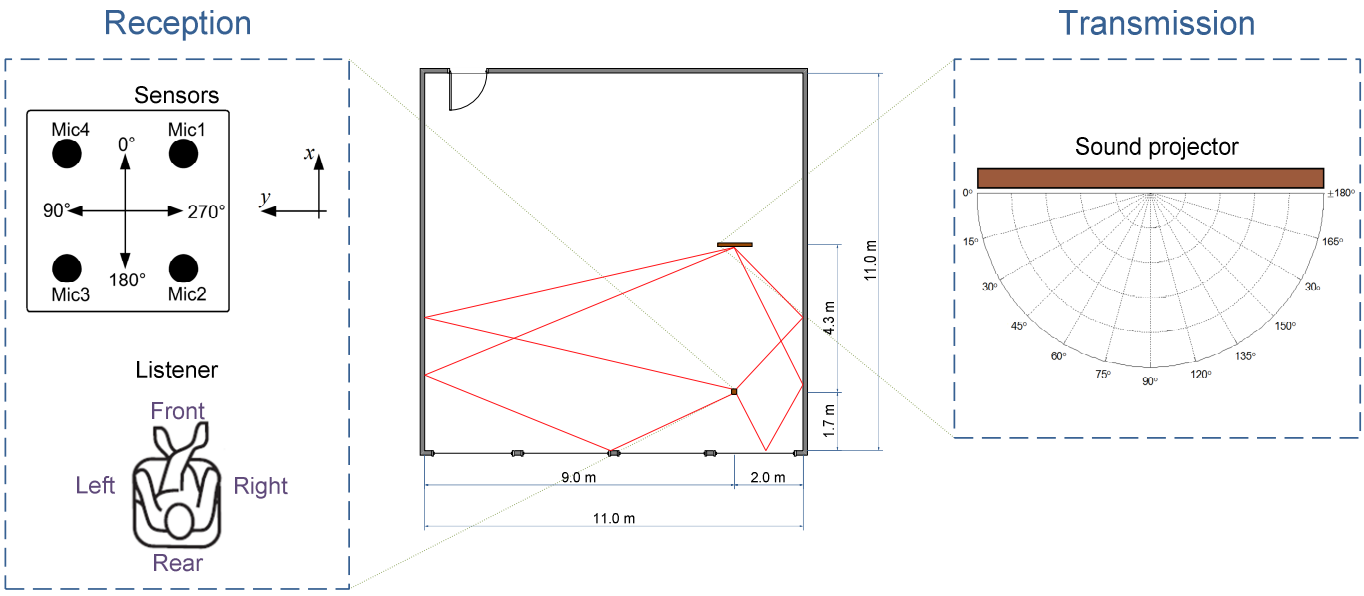



**Figure 2.1.** Floor plan of the laboratory room and coordinate system of the transmitter and receptor sources. The transmission source is the loudspeaker array. At the position of the reception, depending on which phase is running (sounding or projector phase) the sensors or the listener are located. The red lines show the possible path (from the sound projector to the position of the listener) covered by the sound waves for the four described channels (front right and left, rear right and left).

## 2.3 System Components

The hardware used consists of a linear array of 30 loudspeakers, a planar matrix of 4 microphones, the SAM-30 box, a DVD-player, a laptop PC, wires, 4 micro-phone amplifiers and a power supply.

The main components are placed as follows (Figure 2.2): the loudspeaker array is connected to the output channels of the SAM-30 box; the microphones and the Home cinema are connected to a switch box which is also connected to the 4 audio input channels of the DSP board; and, finally, the laptop PC is linked to the DSP using a USB connector. The switch box permits the user select the input signal for the DSP depending on the running state in each moment.

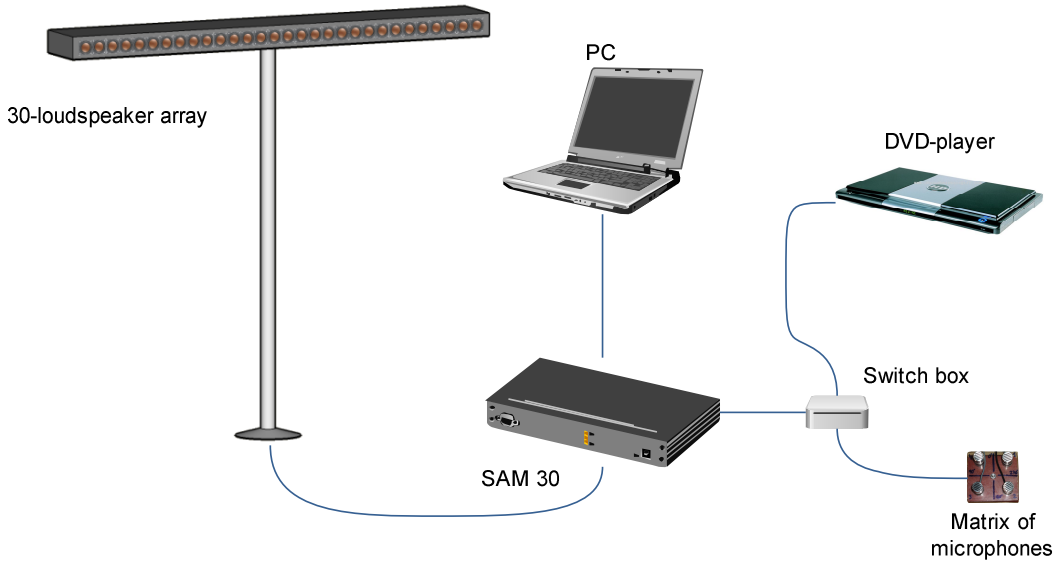



**Figure 2.2.** Diagram of system components.

### 2.3.1 30-Loudspeaker Array

The loudspeaker array consists of 30 loudspeakers model VECO 28KC08-1-A placed side by side in a row with a gap of 3.25 *cm* between each one of them. The total dimension of the array is approximately 1 *m* length. The speakers are circular with a diameter of 2.5 *cm*. Their frequency range is from 150 *Hz* to 20000 *Hz*.

### 2.3.2 SAM-30

The SAM-30 is a complex box which is composed by a DSP board (Texas Instruments 6713), 30 small amplifiers (Analog Devices SSM2211), 30 D/A converters (Analog Devices AD7390), a power supply (ASTEC LPS 60) and all the other needed components to manage the loudspeaker array. It has four input channels, a USB connection and 30 audio output channels connected directly to the loudspeaker array.

The board used is a TMS320C6713 DSK with a floating-point processor from Texas Instruments operating at 225 *MHz*. A DSK_AUDIO4 daughtercard [17], which provides four synchronized, 16-bit A/D and D/A channels, has been mounted on the DSP. Based on the PCM3794 codec, it can sample at 48 *kHz* using its onboard oscillator in default operating mode.

**Figure 2.3.** Picture of the 30-loudspeaker array.


**Figure 2.4.** Picture of the DVD-player and the SAM-30 which contains the DSP.

### 2.3.3 Microphone System

Four omni-directional microphones (AOI Electronics, type ECM-1025) form the record system. Their frequency range response is: 50-16000 *Hz*. They are set in a square with a gap of 4 *cm* between each side, see Figure 2.5.

### 2.3.4 DVD-player

The selected equipment is a Centrum DVD-player, model Triton 410. It can perform a 5.1-channel setup.

Since the DSP card only has four input channels, only the front left/right and the rear left/right channels are used. Thus, the configuration performed by the whole system is a 4.0-channel setup.

**Figure 2.5.** Matrix of microphones.

## 2.4 Running of the System

Subsection 1.4 introduced a brief explanation of the model of the system from a general point of view. In this subsection, the process which sets up the correct configuration of the four channels is described step by step.

The running of the system consists of three phases (sounding, graphics and projector phase), see subsection 1.4, each of which involves several steps (see Figure 2.6):

- **Sounding phase:** The purpose of this phase is to obtain the parameters necessary to configure the four beamformers used in the projector phase. In this phase, the microphones must be connected to the input channels of the DSP.
    - **Beamforming.** Once the system is initialized, the loudspeaker array begins to steer the signal towards a number of specific directions.
    - **Reception of the signals.** For each direction of transmission, the matrix of microphones obtains the sound coming from the different reflections in the room. These signals are stored in the DSP for subsequent analyses.
    - **TDOA and DOA estimation.** After having recorded and stored the signals, the time delay of arrival of signals received from the microphone array can be estimated and thus the direction of arrival is estimated. The obtained DOA is

stored in the sounding table. The sounding table contains the direction of transmission (at the sound projector), the direction of arrival (at the microphones), the sound power received from the DOA and the sound power received from an angle of arrival of 0º (power of direct sound).

- **Decision of the parameters for each channel.** Once the table is filled in with all values (direction of arrival, direction of transmission, power of the DOA and power of direct sound), it is decided which transmit direction and power should be used for each of the four beamformed channels (i.e. front-left, front-right, rear-left, and rear-right). For each channel, which corresponds to a specific angle of arrival, the parameters which maximize the relation between sound power from DOA and power of direct sound are selected.

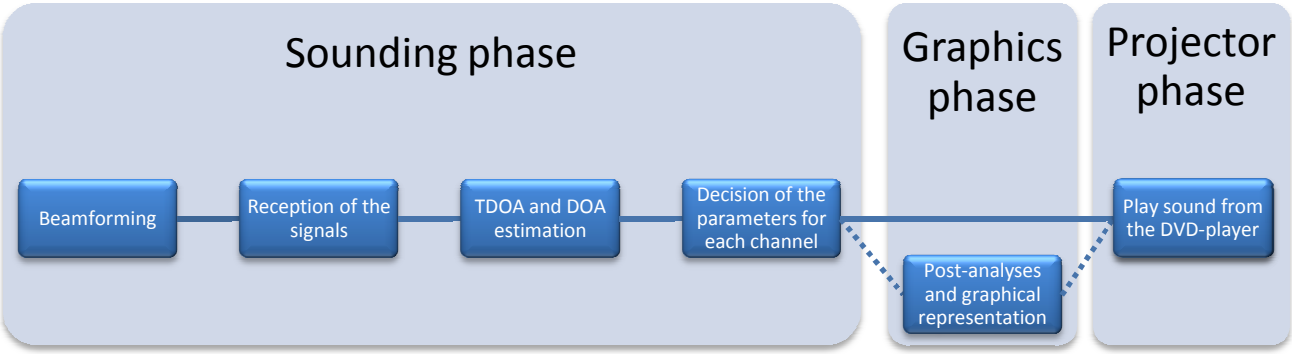



**Figure 2.6.** Block diagram of the system.

- **Graphics phase:** This phase is optional and is not required in order to run the projector phase. Its purpose is to graphically show the results obtained from the sounding phase.
  - **Post-analyses and graphical representation in MATLAB.** In order to run this phase, the samples recorded on the microphones must be loaded to MATLAB from the SAM-30. The post-analyses include both transmission and reception:
    - The beam patterns of the loudspeaker array are shown for each direction of transmission.
    - The received and transmitted signals are shown in frequency and delay domain.
    - Two TDOA analyses (see subsection 3.3) can be run from the received and transmitted signals. From these analyses, the DOA and power of the received signal can be estimated. The first analysis is already performed in the SAM-30, but MATLAB allows the user to graphically see the results. The second analysis is only implemented in MATLAB, but it also obtains the required parameters for the projector phase.

- **Projector phase:** Once the setup has finished, this configuration is used.
    - **Play sound from the DVD-player.** Finally, the system is ready to work as a Home Cinema. In this case, the matrix of microphones is disconnected from the SAM-30 and the DVD-player is connected instead. Then, it plays the signals aimed for each channel (front right/left, and rear right/left channels) with the parameters selected in the previous phase. In order to better enjoy the surround sound effect, the listener should sit at the same position where the sensors were located during the sounding phase.

# Chapter 3

# Methodology

This section describes the procedures and the theoretical analysis developed. The transmission and reception methods as well as the technical specifications of the signal to transmit are explained.

## 3.1 Signal Model

In a surrounding acoustic environment, two situations can be considered: the free-field environment where each sensor receives only the direct-path signal, and reverberant environments where each sensor may receive a large number of reflected signals in addition to direct path. Since the ultimate goal of the sound projector is to play sound in a room, only the second case (reverberant environment) applies here.

Moreover, for this situation, we must differentiate between the single-source case and the multiple-source scenario. During the sounding phase and assuming the worst case, interfering sounds, apart from the sounding signal generated by the loudspeaker array, could appear (human voices, unexpected and punctual noises…). Thus, a multiple-source reverberant model is taken into account.

### 3.1.1 Transmission

For each channel, the transmitted signals by the loudspeaker array are

$$s_i[n] = a \cdot x[n - D_i] \qquad i = 1...30 \tag{3.1}$$

where $x[n]$ is the signal which is going to transmit, $a$, which ranges between 0 and 1, is the attenuation aimed for the current channel and $D_i$ ($i$=1…$N$) are the steering delays for the speaker $i$. For $N$ elements of the array, the values of the signal can be written as an $N \times 1$ vector denoted as

$$\mathbf{S}[n] = \begin{bmatrix} s_1[n] \\ s_2[n] \\ \vdots \\ s_N[n] \end{bmatrix} \qquad \mathbf{X}_D[n] = \begin{bmatrix} x[n-D_1] \\ x[n-D_2] \\ \vdots \\ x[n-D_N] \end{bmatrix} \tag{3.2}$$

$$\mathbf{S}[n] = a \cdot \mathbf{X}_D[n] \tag{3.3}$$

During the projector phase, the signals aimed for the four channels (front right/left and rear right/left) are superimposed. Therefore, the vector can be rewritten as

$$\mathbf{S}[n] = \sum_{l=1}^{4} a_l \cdot \mathbf{X}_{Dl}[n-T_l] = \sum_{l=1}^{4} a_l \cdot \begin{bmatrix} x_l[n-T_l-D_{1l}] \\ x_l[n-T_l-D_{2l}] \\ \vdots \\ x_l[n-T_l-D_{Nl}] \end{bmatrix} \tag{3.4}$$

where $T_l$ ($l$=1…4) are the transmit delays (in samples) aimed for each channel. These delays are used to synchronize the four signals at the position of the listener.

### 3.1.2 Radiated Fields by Arrays

In the case which it is dealing with, the radiation pattern of a single radiating element does not meet the requirements for this certain application, namely: in transmission, control the beam steering. In order to obtain the desired radiation pattern, single radiating elements are assembled in one-dimensional and two-dimensional arrays.

For this purpose, the sound projector is composed by an array of 30 loudspeakers equidistantly spaced (one-dimensional array) and the matrix of microphones consists of a square of 4 sensors (two-dimensional array). This microphone constellation was selected due to the need to find the DOA from the range of 0 to 360 degrees. For the analysis, the matrix of sensors can be decomposed in pairs of microphones with a total of six one-dimensional arrays.

An array of loudspeakers or sensors is a group of $N$ equal elements which transmit or receive sound simultaneously. The radiation pattern of the array is obtained such as the superposition of the radiated fields from each of the elements, whereas in reception the received signal is a linear combination of the picked up signals by each sensor. If the weights and phases of the linear combination in reception are equal to the ones in transmission, their radiation patterns are equals by reciprocity.

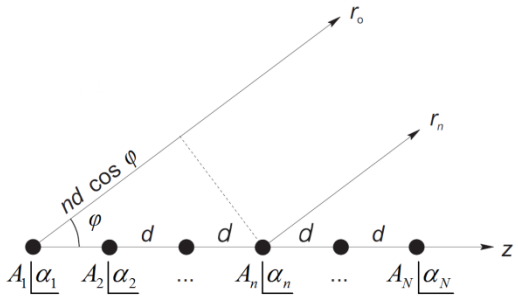



**Figure 3.1.** Uniform linear array of *N* elements.

In this subsection calculations are presented for a one-dimensional, line array consisting of *N* radiating elements with constant spacing *d* between the elements, as shown in Figure 3.1.

In these calculations far-field approach shall be used, meaning that the observation point is located in the far field of the array, so all the paths between the radiating elements and the observation point are considered to be parallel. Furthermore, all path lengths are nearly the same, so there are no significant differences in signal attenuation during propagation due to the differences in path lengths.

The radiation pattern of the array depends on the basic element (loudspeaker or sensor) and the array factor (AF) separately. In order to study the influence of the parameters of the array, independently of the basic element, the array factor is analyzed, which will correspond to the radiation pattern of an array composed by isotropic elements.

The array factor considers the superposition of the *N* sound waves generated by the *N* elements. It only depends on the distance between the elements of the array, the amplitudes and the operating frequency, and it can be calculated as [2]:

$$AF(\varphi) = A_1 + A_2 e^{j\psi(\varphi)} + A_2 e^{j2\cdot\psi(\varphi)} + \ldots + A_N e^{j(N-1)\cdot\psi(\varphi)}$$

$$= \sum_{n=1}^{N} A_n \, e^{j(n-1)\cdot\psi(\varphi)} \tag{3.5}$$

$$\textit{where} \qquad \psi(\varphi) = kd \cdot \cos(\varphi) + \beta$$

The amplitudes of excitation signals supplying the elements are $A_n$, where $n = (1,2,\ldots,N)$, and the phase of excitation signal is $\beta$. $k = 2\pi/\lambda$ is the wave number and $\lambda$ is the wavelength. $\varphi$ is defined as the angle between the positive *z*-axis and the observed direction of the radiation.

Depending on the distribution of amplitudes, different types of linear arrays are defined (uniform, triangular, binomial…). For this application, the uniform distribution is desired. It has the smallest beamwidth, but this comes at the cost of larger sidelobes. Thus, in a uniform linear

array (ULA) the excitation amplitudes are set to one, $A_n$ = 1. Now, equation (3.5) can be simplified as:

$$AF(\varphi) = \sum_{n=1}^{N} e^{j(n-1)\cdot\psi(\varphi)} = e^{j\frac{N-1}{2}\psi(\varphi)} \cdot \frac{\sin\left(\frac{N\cdot\psi(\varphi)}{2}\right)}{\sin\left(\frac{\psi(\varphi)}{2}\right)} \tag{3.6}$$

### 3.1.3 Reception

Assuming that the loudspeaker array is in the sensors' far field, the source radiates a plane wave having the waveform $s[n]$ that propagates through the medium-air. The normal to the wavefront makes an angle $\theta$ with the reference system defined for the sensors (see subsection 2.2), and the signal received at each microphone is a time delayed/advanced version of the signal at a reference sensor (microphone 1).

The selected model assumes a single true source and multiple interfering sources in a reverberant environment. The received signal by each microphone can be expressed as

$$m_l[n] = s[n] * h_l[n] + \sum_{i=1}^{I} t_i[n] * h_{il}[n] + w_l[n] \qquad l = 1\ldots4 \tag{3.7}$$

where $s[n]$ is the known transmitted signal by the loudspeaker array, $t_i[n]$ ($i$=1,2,…,$I$) are the unknown and interfering source signals, which are assumed to be independent with $s[n]$ and with each other. $w_l[n]$ is an additive noise signal at the $l$-th sensor, which is assumed to be uncorrelated with not only all the source signals (both $s[n]$ and $t_i[n]$) but also with the noise observed at other microphones. $h_{il}[n]$ is the channel impulse response from the source $i$ to microphone $l$, which it can be modeled as

$$h_{il}[n] = \sum_{j=1}^{J_i} b_j \cdot \delta\left[n - K_j - D_{m_l m_1}(\theta_j)\right] \tag{3.8}$$

where $J_i$ is the length of the $i$-th channel response, $b_j$ ($j$=1,2,…,$J_i$), which range between 0 and 1, are the attenuation factors due to propagation effects, $K_j$ is the $j$-th propagation delay from the source $i$ to microphone 1 (reference sensor), and $D_{m_l m_1}(\theta_j)$ is the TDOA (also called relative delay) between sensors 1 and $l$ in function of the angle of arrival $\theta_j$.

A simplified model from the general case is shown in equation (3.7) .

$$\mathbf{M}[n]=\begin{bmatrix} m_1[n] \\ \vdots \\ m_4[n] \end{bmatrix}=\begin{bmatrix} b_1\cdot s[n-K_1]+b_2\cdot s[n-K_2] \\ \vdots \\ b_1\cdot s\left[n-K_1-D_{m_4m_1}(\theta_1)\right]+b_2\cdot s\left[n-K_2-D_{m_4m_1}(\theta_2)\right] \end{bmatrix}+ \\ +\begin{bmatrix} b_3\cdot t[n-K_3] \\ \vdots \\ b_3\cdot t\left[n-K_3-D_{m_4m_1}(\theta_3)\right] \end{bmatrix}+\begin{bmatrix} w_1[n] \\ \vdots \\ w_4[n] \end{bmatrix} \tag{3.9}$$

In this equation, a two-coefficient channel impulse response affects the beamsound projector signal and one interfering signal respectively.

## 3.2 Beamforming

Digital time-delay beamforming is used to steer the sound towards a specific direction. The direction-of-transmission resolution is directly related to sampling intervals. In fact, the smaller the sampling intervals the better the direction-of-transmission resolution and the larger the number of natural beams. In a digital time-delay beamformer based on DSP the sampling intervals are primarily restricted by the sampling frequency of the DSP, which is 48 kHz in this case. However, interpolation techniques could be used to choose the desired sampling intervals.

Thus, the number of directions that the main beam can be steered is restricted by the sampling intervals. Assuming that the main beam must be steered at the angle of transmission then the signals which feed two sequential elements must be shifted by

$$\tau_s=\frac{d\cdot\cos(\varphi)}{v_s} \tag{3.10}$$

where $v_s$ is the speed of sound. Since the signals are stored and treated in the DSP in a digital domain, equation (3.10) can be rewritten as

$$\tau_s=\frac{d\cdot\cos(\varphi)}{v_s}=\frac{D(\varphi)}{f_s} \qquad \rightarrow \qquad \cos(\varphi)=\frac{D(\varphi)\cdot v_s}{d\cdot f_s} \tag{3.11}$$

where $D(\varphi)$ is the steering delay in samples and $f_s$ is the sampling frequency.

Thus, for the coordinate system used at the loudspeaker array, the maximum and the minimum delay are obtained when $\varphi_{\min} = 0^o$ and $\varphi_{\max} = 180^o$ respectively:

$$f_s = 48000\ Hz, \quad d = 0.0325\ m, \quad v_s = 340\ m/s$$

$$\varphi_{\min} = 0^o \qquad \rightarrow \qquad D_{\max} = D(0) = \frac{d \cdot f_s}{v_s} = 4.59\ samples \tag{3.12}$$

$$\varphi_{\max} = 180^o \qquad \rightarrow \qquad D_{\min} = D(180) = -\frac{d \cdot f_s}{v_s} = -4.59\ samples \tag{3.13}$$

Therefore, the beam can be steered towards 9 directions, using delays from -4 to 4 samples. In order to increment the number of angles of transmission an interpolation method has to be used.

In the present case, interpolation by 2 was chosen. Interpolation by 4 was attempted, but the use of low-pass filtering (see subsection 5.2.3) and the increase of the processing cycles made the DSP fail the real-time condition. The fact that an interpolation by 2 is used means that 19 different directions of transmission are possible. Table 3.1 illustrates both steering delays and subsequent angles used in the DSP program to direct the beam. The radiation patterns of the directions of transmission are shown in Appendix A .

| Step of the sweep | *D*[*samples*] | *φ*[*degrees*] |
|---|---|---|
| **1** | 4.5 | 11.3° |
| **2** | 4 | 29.3° |
| **3** | 3.5 | 40.3° |
| **4** | 3 | 49.2° |
| **5** | 2.5 | 57.0° |
| **6** | 2 | 64.2° |
| **7** | 1.5 | 70.9° |
| **8** | 1 | 77.4° |
| **9** | 0.5 | 83.7° |
| **10** | 0 | 90.0° |
| **11** | -0.5 | 96.3° |
| **12** | -1 | 102.6° |
| **13** | -1.5 | 109.1° |
| **14** | -2 | 115.8° |
| **15** | -2.5 | 123.0° |
| **16** | -3 | 130.8° |
| **17** | -3.5 | 139.7° |
| **18** | -4 | 150.7° |
| **19** | -4.5 | 168.7° |

**Table 3.1.** List of steering delays and angles, in relation with the coordinate system of the sound projector, in function of the step of the sweep when an interpolation by 2 is applied.

This means that if the value of the signal at the first element is equal to $s_1[n] = \alpha \cdot x[n]$ then the value of the signal at the second element is equal to $s_2[n] = \alpha \cdot x[n-D]$. Expression (3.14) shows the output signal for speaker *i*.

$$s_i[n] = a \cdot x\left[n - (i-1) \cdot D\right] \qquad i = 1...30 \tag{3.14}$$

Bearing in mind the transmission signal model (subsection 3.1.1), the transmitted signal can be rewritten as

$$\mathbf{S}[n] = a \cdot \mathbf{X}_D[n] = a \cdot \begin{bmatrix} x[n] \\ x[n-D] \\ x[n-2D] \\ \vdots \\ x[n-(N-1) \cdot D] \end{bmatrix} \tag{3.15}$$

During the projector phase, the signals aimed for different channels (i.e. rear-left, rear-right, front-left, front-right) are super-positioned at the input of the array. Thus, the signal model is the following.

$$\mathbf{S}[n] = \sum_{l=1}^{4} a_l \cdot \mathbf{X}_{Dl}[n - T_l] = \sum_{l=1}^{4} a_l \cdot \begin{bmatrix} x_l[n - T_l] \\ x_l[n - T_l - D_l] \\ x_l[n - T_l - 2D_l] \\ \vdots \\ x_l[n - T_l - (N-1) \cdot D_l] \end{bmatrix} \tag{3.16}$$

## 3.3 TDOA Estimation

The TDOA estimation problem is concerned with the measurement of time difference between the signals received at different microphones.

When a signal is measured by two spatially separated microphones, some time shift between the two collected recordings is expected. This lag time depends on the DOA of the signal (the angle between the microphone plane and the sound wave).

Two wideband techniques have been implemented to estimate the time delays of arrival. They belong to the family of the Generalized Cross-Correlation (GCC) methods proposed by Knapp and Carter [11]. These two methods work properly with uncorrelated wideband signals.

The GCC method only considers two microphones. Therefore, the TDOA estimate is obtained by finding the number of lag samples that maximize the estimated cross-correlation function between the filtered signals of the microphone outputs:

$$\hat{D}_{m_i m_j} = \arg\max_p \left( \hat{r}^{CCG}_{m_i m_j} [p] \right) \tag{3.17}$$

where

$$\begin{aligned} \hat{r}^{CCG}_{m_i m_j} [p] &= F^{-1} \left[ \hat{\Psi}_{m_i m_j} (f) \right] \\ &= \int_{-\infty}^{\infty} \hat{\Psi}_{m_i m_j} (f) \cdot e^{i 2\pi f p} \, df \\ &= \int_{-\infty}^{\infty} v(f) \cdot \hat{G}_{m_i m_j} (f) \cdot e^{i 2\pi f p} \, df \end{aligned} \tag{3.18}$$

is the cross-correlation function for the GCC method,

$$\hat{G}_{m_i m_j} (f) = M_i (f) \cdot M_j (f) \tag{3.19}$$

is the estimated cross-spectrum of the received signals and $v(f)$ denotes the general frequency weighting.

There are many different choices of the frequency-domain weighting function, leading to a variety of different GCC methods. The main important of them are the following [10] [11]:

- **Classical Cross-Correlation.** It obtains the TDOA from the usual cross-correlation function. In this case the weighting function is $v(f) = 1$.
- **Smoothed Coherence Transform (CGG-SCOT).** In order to overcome the impact of fluctuating levels of the speech source signal on TDOA estimation, an effective way is to pre-whiten the microphone outputs before their cross-spectrum is computed. The weighting function used is

$$v(f) = \frac{1}{\sqrt{\hat{G}_{m_i} (f) \cdot \hat{G}_{m_j} (f)}}$$

- **Phase Transform (GCC-PHAT).** This algorithm combats the room reverberation effects due to multi-path very well. It pre-whitens the signals using the information included in its phase. The weighting function used is

$$v(f) = \frac{1}{\left| \hat{G}_{m_i m_j} (f) \right|}$$

The techniques explained below are derived from the Classical Cross-Correlation method, since it is wanted to keep the reverberation effects and analyze their behavior. The first technique implements the Classical Cross-Correlation method, but the second introduces some modifications and improvements with the adding of the transmitted signal in the calculations.

### 3.3.1 Maximum Value of TDOA

The maximum value of the TDOA between two sensors is given by the distance between those sensors. Since the matrix of microphones consists of a square of 4 microphones, it can be decomposed into six 2-element ULAs. Each side of the square is 4 cm long and the two diagonals are 5.66 cm.

Due to the fact that the transmission and reception theory of ULAs follows the same theoretical pattern, equation (3.11) can be reused in this context.

$$\cos(\theta) = \frac{D_{m_i m_j}(\theta) \cdot v_s}{d_{m_i m_j} \cdot f_s} \tag{3.20}$$

where $d_{m_i m_j}$ is the distance between sensor *i* and *j*. The maximum value of the TDOA is given when $\theta = 0^o$.

$$D_{m_i m_j \max} = D_{m_i m_j}(0) = \frac{d_{m_i m_j} \cdot f_s}{v_s} \tag{3.21}$$

Thus, this distance is converted to samples in order to determine its maximum delay:

- For each side (e.g. microphone 1 and 4):

$$d_{m_1 m_4} = 4\,cm$$

$$D_{m_1 m_4 \max} = \frac{d_{m_1 m_4} \cdot f_s}{v_s} = 5.65\,samples \quad \rightarrow \quad 6\,samples \tag{3.22}$$

- For the diagonals (e.g. microphone 1 and 3):

$$d_{m_1 m_3} = 5.66\,cm$$

$$D_{m_1 m_3 \max} = \frac{d_{m_1 m_3} \cdot f_s}{v_s} = 7.99\,samples \quad \rightarrow \quad 8\,samples \tag{3.23}$$

Now, assuming the case in which a signal without any reflection is recorded, it can be stated that this signal will be the same for each microphone but a delay of some samples. Then, the propagation delay between two different sensors varies depending on the DOA, but this number is always ranged in the interval [-6, 6] for pairs of microphones corresponding to each side and [-8, 8] for pairs corresponding to the diagonals.

### 3.3.2 Technique 1: Classical Cross-Correlation (CC)

This is the most simple and straightforward method of TDOA estimation. However, its performance is often affected by many factors such as signal self correlation, reverberation, interfering signals, etc.

Within the context of this project, reverberation effects are important in order to see which paths are followed by the sound waves. One of the main drawbacks of this method is the error produced by interfering signals since this method is not capable to delete it and the measurements will include these signals. Thus, the results can be distorted and consequently useless. Therefore, the first requirement in order to use this method is to have a silent environment to avoid any annoying sound.

#### 3.3.2.1 Search for the TDOA between Signals

This technique is used to develop the TDOA analysis implemented in the SAM-30 and in MATLAB. It obtains the six TDOA corresponding to the six pairs of microphones. Once those delays have been obtained the DOA can be calculated as shown in subsection 3.4.1.

The development of this technique can be followed in the spectral or temporal analysis. Herein, we implemented this method on the DSP in the time domain. In the following, the procedure is shown in the discrete time.

This technique is based on the cross-correlation of the recorded signals between pairs of microphones. Then, in this case, six pairs of different cross-correlations are estimated.

$$\hat{r}_{m_i m_j}[p] = m_i[p] * m_j^*[-p] \tag{3.24}$$

$$\hat{r}_{m_i m_j}[p] = \begin{cases} \frac{1}{L}\sum_{n=0}^{L-p-1} m_i[n+p]\cdot m_j^*[n] & p \geq 0 \\ \hat{r}_{m_i m_j}[-p] & p < 0 \end{cases} \tag{3.25}$$

where $L$ is the total length in samples of $m[n]$. This time-averaged estimate is a biased estimator with low estimate variance and asymptotically unbiased [4][10].

As it has been explained in subsection 3.3.1, the maximum delay between two recorded signals by the microphones is fixed. Since the sound signals are real, equation (3.25) can be simplified as follows:

$$\hat{r}_{m_i m_j}[p] = \begin{cases} \frac{1}{L}\sum_{n=0}^{L-p-1} m_i[n+p]\cdot m_j[n] & p \in \left[0, D_{m_i m_j \max}\right] \\ \hat{r}_{m_i m_j}[-p] & p \in \left[-D_{m_i m_j \max}, 0\right) \end{cases} \tag{3.26}$$

Where $D_{m_i m_j \max}$ is 6 or 8 samples depending on the chosen pair. Then, the value which maximizes $\hat{r}_{m_i m_j}[p]$ gives the relative delay between signals. Bearing in mind equation (3.17).

$$\hat{D}_{m_i m_j} = \arg\max_{p}\left(\hat{r}_{m_i m_j}[p]\right)$$

This manner, the six delays are obtained and the DOA can be estimated from the range of 0º to 360º (see subsection 3.4.1).

$$\hat{\theta} = f\left(\hat{D}_{m_2 m_1}, \hat{D}_{m_3 m_4}, \hat{D}_{m_1 m_4}, \hat{D}_{m_2 m_3}, \hat{D}_{m_2 m_4}, \hat{D}_{m_1 m_3}\right) \qquad \hat{\theta} \in [0, 360)$$

#### 3.3.2.2 Estimated Power in Function of the Angle

This technique is used during the graphic phase. It obtains a polar graphic of the estimated power in function of the DOA (from all the directions of the plane $\theta \in [0,2\pi)$) in MATLAB.

The four recorded signals are delayed a number of samples according to a fixed angle of arrival. Then they are added and the power in function of the DOA is estimated.

$$m_T[n,\theta] = \frac{1}{4}\left(m_1[n] + \sum_{i=2}^{4} m_i\left[n - D_{m_i m_1}(\theta)\right]\right) \tag{3.27}$$

where $D_{m_i m_1}(\theta)$ is the delay from sensor *i* to 1 in function of a specific angle. Thus, the received sound power is obtained as a function of the angle.

$$\hat{P}_T[\theta] = \frac{1}{L}\sum_{n=0}^{L-1} m_T^2[n,\theta] \tag{3.28}$$

#### 3.3.2.3 Estimated Instantaneous Power in Function of Delay and Angle

This technique is used during the graphic phase. It obtains a 3D graphic of the estimated instantaneous power in function of the time and the DOA (from all the directions of the plane $\theta \in [0,2\pi)$ ) in MATLAB.

The four recorded signals are delayed a number of samples according to a fixed angle of arrival. Then they are added and the instantaneous power in function of the DOA is estimated (see equation (3.27)). Thus, the received sound power is obtained as a function of delay and angle.

$$\hat{P}_i[n,\theta] = m_T^2[n,\theta] \tag{3.29}$$

### 3.3.3 Technique 2: Cross-Correlation with the Source Signal (CSS)

This method is also based on the cross-correlation of the recorded signals but it takes advantage from the fact that the sent signal is known. Thus, it modifies slightly the procedure of the GCC methods, which only use the received signals. This way, this technique is capable of solving the drawback which appeared in the previous method: the lack of robustness against external signals. This technique It is only developed in MATLAB.

The implementation of the CSS technique is much more complex than CC technique and it increments the processing time considerably.

#### 3.3.3.1 Estimated Cross-Correlation in Function of Delay and Angle

This technique is used during the graphic phase. It obtains a 3D graphic of the cross-correlation between the received and the transmitted signals as function of propagation delay and angle (from all the directions of the plane $\theta \in [0,2\pi)$ ).

This method uses the delay-and-sum analysis in order to obtain the cross-correlation between the recorded and the sent signals depending on the DOA from a range of 0 to 360 degrees (0 to $2\pi$, in radians).

The four received signals are delayed a number of samples according to a fixed angle of arrival and they are added as

$$m_T[n,\theta] = \frac{1}{4}\left( m_1[n] + \sum_{i=2}^{4} m_i\left[ n - D_{m_i m_1}(\theta) \right] \right)$$

Then, the combined signal is cross-correlated with the transmitted signal. In this case, the implementation of this algorithm is faster in the frequency domain, so the Fourier Transform is used.

$$\begin{aligned} \hat{r}_{m_T s}[p,\theta] &= F^{-1}\left[ \hat{G}_{m_T s}(f,\theta) \right] \\ &= \int_{-\infty}^{\infty} \hat{G}_{m_T s}(f,\theta) \cdot e^{i2\pi fp}\, df \end{aligned} \qquad p \in \left[ -(L-1), L-1 \right], \quad \theta \in [0, 2\pi) \tag{3.30}$$

where $L$ is the length of the signals and

$$\begin{aligned} \hat{G}_{m_T s}(f,\theta) &= M_T(f,\theta) \cdot S^*(f) \\ &= \frac{1}{4}\left[ M_1(f) + M_2(f,\theta) + M_3(f,\theta) + M_4(f,\theta) \right] \cdot S^*(f) \end{aligned} \tag{3.31}$$

is the estimated cross-spectrum between $m_T[n, \theta]$ and the source signal $s[n]$. As for negative values of $p$ the estimated cross-correlation refers to samples before beginning the transmission and reception, which does not make any sense in this context, the limits of equation (3.30) can be rewritten as follows.

$$p \in [0, L-1], \quad \theta \in [0, 2\pi) \tag{3.32}$$

The values of $p$ and $\theta$ which maximize $\hat{r}_{m_T s}[p, \theta]$ show the propagation delay and the DOA of the wavefront that arrived more powerfully.

$$\left(\hat{K}, \hat{\theta}\right) = \arg\max_{p,\theta} \left(\hat{r}_{m_T s}[p, \theta]\right) \tag{3.33}$$

### 3.3.3.2 Estimated Power in Function of the Angle

This technique is used during the graphic phase. It obtains a polar graphic of the estimated power in function of the DOA (from all the directions of the plane $\theta \in [0,2\pi)$) in MATLAB.

The estimated cross-correlation $\hat{r}_{m_T s}[p, \theta]$, obtained from the previous subsection, is also used in this technique. The Otsu method [18] is applied in order to determine the threshold between the background noise and the peaks, which indicate the arrival of new sound beams at the position of the microphones. The Otsu method is one of the most referenced thresholding methods[19].

First, the Otsu's threshold is calculated as a function of $\hat{r}_{m_T s}[p, \theta]$. Once it has been obtained, the estimated power in function of the angle can be obtained as

$$\hat{r}_O[p, \theta] \begin{cases} \hat{r}_{m_T s}[p, \theta] & \hat{r}_{m_T s}[p, \theta] > t_O \\ 0 & \hat{r}_{m_T s}[p, \theta] < t_O \end{cases} \tag{3.34}$$

$$\hat{P}_{m_T s}[\theta] = \frac{1}{L} \sum_{p=0}^{L-1} \hat{r}_O[p, \theta] \tag{3.35}$$

where $t_O$ is the Otsu's threshold and $\hat{P}_{m_T s}[\theta]$ is the estimated power in function of the angle.

### 3.3.4 Robustness of Both Methods under Interfering and Noise Signals

In this subsection, the robustness of both methods is discussed from a theoretical point of view. The signal model is composed by the reception of two scatters of the transmitted signal, an interfering signal and additive noise (see subsection 3.1.3).

Bearing in mind equation (3.9), the received signals at the microphones are

$$\mathbf{M}[n]=\begin{bmatrix} m_1[n] \\ \vdots \\ m_4[n] \end{bmatrix}=\begin{bmatrix} b_1\cdot s[n-K_1]+b_2\cdot s[n-K_2] \\ \vdots \\ b_1\cdot s\left[n-K_1-D_{m_4m_1}(\theta_1)\right]+b_2\cdot s\left[n-K_2-D_{m_4m_1}(\theta_2)\right] \end{bmatrix}+$$
$$+\begin{bmatrix} b_3\cdot t[n-K_3] \\ \vdots \\ b_3\cdot t\left[n-K_3-D_{m_4m_1}(\theta_3)\right] \end{bmatrix}+\begin{bmatrix} w_1[n] \\ \vdots \\ w_4[n] \end{bmatrix}$$

It is assumed that all the source signals are mutually independent, $s[n]$ is uncorrelated with $t[n]$ and $w_i[n]$ ($i$=1,…,4), and $w_i[n]$ is uncorrelated with both the source signals ($s[n]$ and $t[n]$) and the noise observed at other sensors.

#### 3.3.4.1 Robustness of CC Technique

The cross-correlation function of the signals $m_i[n]$ and $m_j[n]$ equals

$$\hat{r}_{m_im_j}[p]=b_1^2\cdot\hat{r}_s\left[p-D_{m_im_j}(\theta_1)\right]+b_2^2\cdot\hat{r}_s\left[p-D_{m_im_j}(\theta_2)\right]+$$
$$+b_3^2\cdot\hat{r}_t\left[p-D_{m_im_j}(\theta_3)\right]+$$
$$+b_1b_2\cdot\hat{r}_s\left[p-(K_1-K_2)-\left(D_{m_im_1}(\theta_1)-D_{m_jm_1}(\theta_2)\right)\right]+$$
$$+b_1b_2\cdot\hat{r}_s\left[p-(K_2-K_1)-\left(D_{m_im_1}(\theta_2)-D_{m_jm_1}(\theta_1)\right)\right] \qquad p\in\left[-D_{m_im_j\max},D_{m_im_j\max}\right]$$

Only the first two components are desired, which are the contribution of two different sound waves coming from the sound projector. The third component is the contribution of the interfering signal, which cannot be removed, and the other two components are caused by the cross-correlation between the two sound waves coming from the sound projector.

In order to avoid the effects of these two last components as much as possible, the maximum value of them should be shifted outside the analyzed region. Thus, the following restriction has to be met.

$$\left|\Delta K+\left(D_{m_i m_1}(\theta_2)-D_{m_j m_1}(\theta_1)\right)\right|>D_{m_1 m_3 \max} \tag{3.36}$$

where $\Delta K=K_2-K_1$. The analysis of the positive case of (3.36) gives the first condition to meet.

$$\Delta K+\left(D_{m_i m_1}(\theta_2)-D_{m_j m_1}(\theta_1)\right)>D_{m_1 m_3 \max}$$

The worst case appears when the delays get the maximum values.

$$\begin{cases} D_{m_i m_1}(\theta_2)=-D_{m_1 m_3 \max} \\ D_{m_j m_1}(\theta_1)=D_{m_1 m_3 \max} \end{cases}$$

Therefore, the maximum lag samples between two consecutive sound waves is given by (3.37).

$$\Delta K+\left(-D_{m_1 m_3 \max}-D_{m_1 m_3 \max}\right)>D_{m_1 m_3 \max}$$

$$\Delta K>3\cdot D_{m_1 m_3 \max} \tag{3.37}$$

The second condition is given by a similar analysis done for the negative case of equation (3.36)

$$\Delta K+\left(D_{m_i m_1}(\theta_2)-D_{m_j m_1}(\theta_1)\right)<-D_{m_1 m_3 \max}$$

In this case, the values which restrict the previous equation the most are the following.

$$\begin{cases} D_{m_i m_1}(\theta_2)=D_{m_1 m_3 \max} \\ D_{m_j m_1}(\theta_1)=-D_{m_1 m_3 \max} \end{cases}$$

Then, the second condition is obtained.

$$\Delta K<-3\cdot D_{m_1 m_3 \max} \tag{3.38}$$

From equations (3.37) and (3.38), the following constraint can be deduced.

$$|\Delta K|>3\cdot D_{m_1 m_3 \max} \tag{3.39}$$

Concluding, the presented method is able to recognize the correct directions of arrival in a multi-path channel if the consecutive scatters of the sound are separated by at least $|\Delta t_{\min}|$ such that

$$|\Delta t_{\min}|=\frac{|\Delta K|}{f_s}=\frac{3\cdot D_{m_1 m_3 \max}}{f_s}=500\mu s \tag{3.40}$$

Moreover, this technique is totally affected by the interfering signals. They cannot be removed and perturb the correct DOA estimation of the signals transmitted by the sound projector.

### 3.3.4.2 Robustness of CSS Technique

The cross-correlation function of the $m_T[n,\theta]$ and $s[n]$ signals equals

$$\hat{r}_{m_T s}[p,\theta] = \frac{b_1}{4}\cdot\left(\hat{r}_s[p-K_1]+\sum_{i=2}^{4}\hat{r}_s\left[p-K_1-\left(D_{m_i m_1}(\theta_1)-D_{m_i m_1}(\theta)\right)\right]\right)+$$
$$+\frac{b_2}{4}\cdot\left(\hat{r}_s[p-K_2]+\sum_{i=2}^{4}\hat{r}_s\left[p-K_2-\left(D_{m_i m_1}(\theta_2)-D_{m_i m_1}(\theta)\right)\right]\right)$$

As can be seen, unlike the behavior observed in the first method, the CSS technique removes the contribution of the interfering signal. It only appears the contribution of the signal transmitted by the loudspeaker array affected by the channel impulse response.

Thus, when the angle of arrival $\theta_1$ coincides with the, i. e. $\theta_1 = \theta$, the cross-correlation has a peak for $p = K_1$, which is the propagation delay

$$\hat{r}_{m_T s}[p,\theta] = b_1\cdot\hat{r}_s[p-K_1]+\frac{b_2}{4}\cdot\left(\hat{r}_s[p-K_2]+\sum_{i=2}^{4}\hat{r}_s\left[p-K_2-\left(D_{m_i m_1}(\theta_2)-D_{m_i m_1}(\theta)\right)\right]\right)$$

# 3.4 DOA Estimation

In this subsection, a method for estimating the angle of the signal incoming from an arbitrary direction from the range of 0 to 360 degrees is described.

Since the previous techniques are developed in terms of relative delays between sensors (TDOAs) and in terms of angle of arrival (DOA), it is necessary to convert them as its opposite form: i.e from TDOAs to DOA, and from DOA to TDOAs.

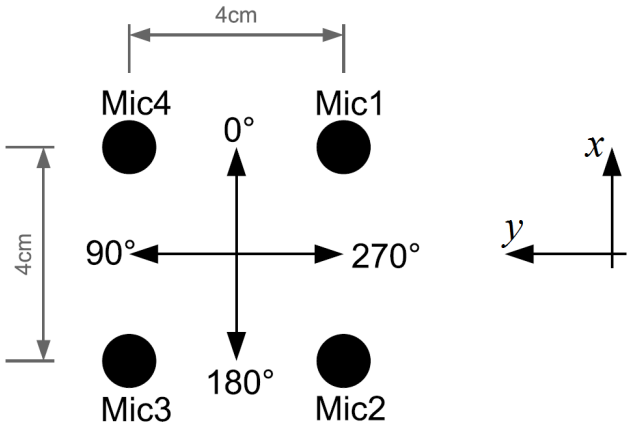


**Figure 3.2.** Relation between the Cartesian and the Polar coordinate systems for the matrix of microphones.

### 3.4.1 Obtaining the DOA from the TDOAs

This conversion is used in the technique explained in subsection 3.3.2.1. It combines the six TDOAs between pairs of microphones in order to find the angle of arrival $\theta$ from the range of 0º to 360º.

$$\theta = f\left(D_{m_2m_1}, D_{m_3m_4}, D_{m_1m_4}, D_{m_2m_3}, D_{m_2m_4}, D_{m_1m_3}\right) \qquad \theta \in [0, 360)$$

Bearing in mind the Cartesian and the Polar coordinate systems defined in subsection 2.2, they can be combined as follows. The delays obtained from pairs of sensors 1-4 and 2-3 pertain to the "*y*" component, the ones obtained from pairs 3-4 and 2-1 pertain to the "*x*" component, and the last group, obtained from pairs 2-4 and 1-3, need to be decomposed on "*x*" and "*y*" components.

$$\theta = \arctan\left(\frac{y_T}{x_T}\right) \qquad where \qquad \begin{cases} y_T = y_{m_1m_4} + y_{m_2m_3} + y_d \\ x_T = x_{m_2m_1} + x_{m_3m_4} + x_d \end{cases} \tag{3.41}$$

$x_{m_im_j}$ and $y_{m_im_j}$ are the "*x*" and "*y*" component in samples of the pair of microphones *i* and *j*, and $x_d$ and $y_d$ are the "*x*" and "*y*" components of the pairs in the diagonals. As explained above, for this reference system, the $y_{m_2m_1}$, $y_{m_3m_4}$, $x_{m_1m_4}$ and $x_{m_2m_3}$ components are zero. We have:

$$\begin{cases} x_{m_2m_1} = D_{m_2m_1} \\ x_{m_3m_4} = D_{m_3m_4} \\ x_{m_1m_4} = 0 \\ x_{m_2m_3} = 0 \end{cases} \qquad \begin{cases} y_{m_2m_1} = 0 \\ y_{m_3m_4} = 0 \\ y_{m_1m_4} = D_{m_1m_4} \\ y_{m_2m_3} = D_{m_2m_3} \end{cases}$$

In order to obtain $x_d$ and $y_d$ from the delay obtained from the pairs of sensors in the diagonals, the following conversion has to be applied.

$$\tan(\alpha) = \frac{D_{m_2m_4}}{D_{m_1m_3}} \tag{3.42}$$

$$\begin{cases} x_d = \dfrac{d_{m_1m_3} \cdot f_s}{v_s} \cdot \cos\left(\alpha - 45^o\right) \\ y_d = \dfrac{d_{m_1m_3} \cdot f_s}{v_s} \cdot \sin\left(\alpha - 45^o\right) \end{cases} \tag{3.43}$$

### 3.4.2 Obtaining the TDOAs from the DOA

This conversion is used in the techniques explained in subsection 3.3.2.3 and 3.3.3. It obtains the six TDOAs between pairs of sensors from a specific angle of arrival $\theta$.

For the defined reference system the delay can be calculated as follows.

$$\begin{cases} D_{m_2m_1} = \dfrac{d_{m_1m_4} \cdot f_s}{v_s} \cdot \cos(\theta) \\ D_{m_3m_4} = D_{m_2m_1} \\ D_{m_1m_4} = \dfrac{d_{m_1m_4} \cdot f_s}{v_s} \cdot \sin(\theta) \\ D_{m_2m_3} = D_{m_1m_4} \\ D_{m_1m_3} = \dfrac{d_{m_1m_3} \cdot f_s}{v_s} \cdot \cos(\theta + 45^o) \\ D_{m_2m_4} = \dfrac{d_{m_1m_3} \cdot f_s}{v_s} \cdot \sin(\theta + 45^o) \end{cases} \tag{3.44}$$

## 3.5 Acoustic Frequency Range

In the previous subsections, it has been explained how to steer or find the direction of the main lobe in an array of loudspeakers and in a matrix of microphones respectively. The fact is that depending on the geometry of these arrays and the frequency of the signal used, the radiation pattern (transmitted and received) may vary adding grating lobes (due to aliasing effects) or losing directivity in the main lobe.

In order to obtain better results, the bandwidth of the transmitted signal must be limited. During the sounding phase, the architecture of the sensors and the sound projector must be taken into account; meanwhile, during the projector phase the sensors are not connected to the DSP and a human listener is set at their position. Therefore, a different bandwidth is applied in each phase.

### 3.5.1 Frequency Range Limited by the Loudspeaker Array

In this subsection, the optimum bandwidth for a transmitted signal is found from the point of view of the sound projector. The conditions which should meet the sent signal are: not to add

spatial aliasing, i. e. to avoid the appearance of alias beams or grating lobes, and to have a good directivity.

The maximum and the minimum frequency are limited by the distance between elements on the loudspeaker array. Bearing in mind equation (3.6):

$$AF(\theta) = e^{j\frac{N-1}{2}\psi(\theta)} \cdot \frac{\sin\left(\frac{N\cdot\psi(\theta)}{2}\right)}{\sin\left(\frac{\psi(\theta)}{2}\right)} \qquad where \qquad \psi(\theta) = kd\cdot\cos(\theta)+\beta$$

Due to the fact that angle $\theta$ only takes real values between 0 and $\pi$, the interval of variation of $\psi(\theta)$ can be determined as

$$\psi(\theta) \in [-kd+\beta, kd+\beta] \tag{3.45}$$

Thus the interval ranged in (3.45) belongs to the radiation pattern and it is called visible region. Its length is $2kd$ and it is centered in $\beta$.

The beamwidth of the major lobe varies with the wavelength, or, in terms of frequency, the higher the frequency, the narrower the beamwidth is. But if the frequency increases too much another main lobe can appear in the visible range.

The aim of this application is to steer the sound in a specific angle, so the addition of this new lobe is not desired. If it appeared, the power of the main lobe would be split between the main and the grating lobe, making it decrease the power transmitted in the desired direction. In the same way, the sound would be transmitted in two different angles.

The worst case is when the sound projector transmits the signal in the end-fire direction ($\theta$ = 0º or $\theta$ = 180º). In this situation, two main lobes appear in the visible range and perturb the correct direction of transmission. The usual analysis of an end-fire array [3] shows that in order to avoid it the following constraint should be met:

$$d < \frac{\lambda}{2} \tag{3.46}$$

Consequently, the maximum frequency should be the following:

$$v_s = 340\,m/s, \quad d = 0.0325\,m$$

$$f_{\max} = \frac{v_s}{\lambda} = \frac{v_s}{2\cdot d} = 5230.8\,Hz \tag{3.47}$$

However, if a deeper analysis is done, this condition does not totally avoid the grating lobe, and it still appears in the frequencies near the maximum.

For large arrays this situation can be solved without reducing the frequency range too much. Since the loudspeaker array is composed by 30 elements, the following analysis applies. It consists of defining the maximum visible range of an end-fire array, as Figure 3.3 shows.

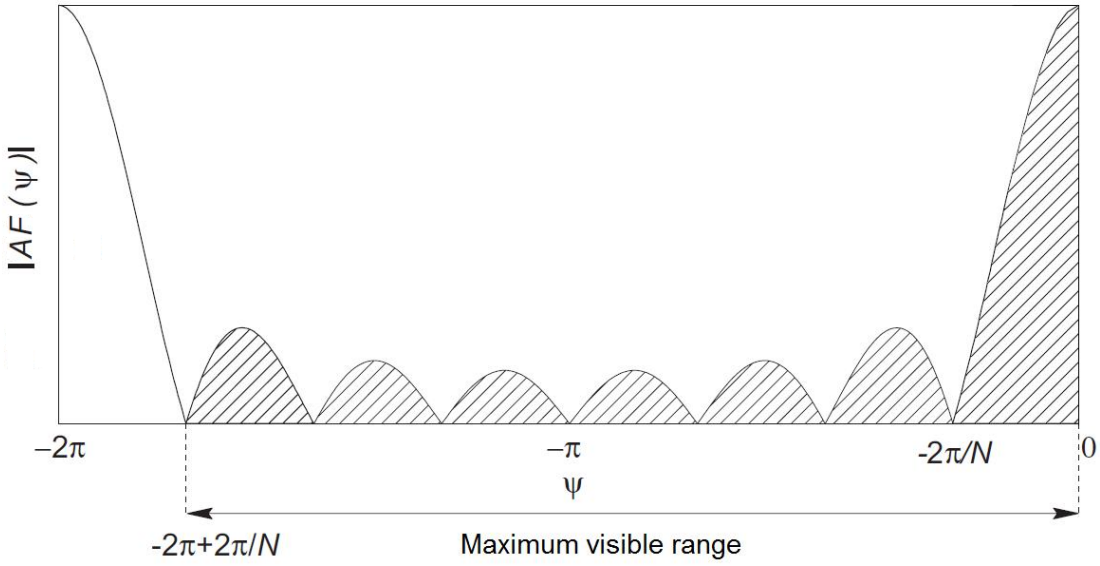



**Figure 3.3.** Maximum visible range of an *N*-element end-fire array ($\theta = 0^{o}$).

Then, the maximum visible range does not include the grating lobe located at the left. Thus, the following condition should be met.

$$2kd < 0 - \left( -2\pi + \frac{2\pi}{N} \right) = 2\pi \cdot \left( \frac{N-1}{N} \right) \tag{3.48}$$

$$k = \frac{2\pi}{\lambda} \quad \rightarrow \quad \lambda > \frac{2d \cdot N}{N-1} \tag{3.49}$$

Therefore, according to (3.49), the new maximum frequency is calculated as

$$N = 30$$

$$\lambda > \frac{2d \cdot N}{N-1} \rightarrow \lambda_{\min} = 0.0672\, m$$

$$f'_{\max} = \frac{v_s}{\lambda_{\min}} = 5056.4\, Hz \tag{3.50}$$

Comparing (3.47) and (3.50) it can be observed that the loss of spectral width is minimal when compared to the advantage of avoiding that lobe. As it has been seen, the maximum frequency which can be used without the appearance of a grating lobe is 5056.4 Hz. For lower frequencies this lobe is outside the edge of the visible range.

In order to determine the minimum frequency a similar analysis has to be done. As explained above, the beamwidth of the main lobe varies with the frequency. In this case: the lower the frequency, the wider the beamwidth is.

The sound projector has useful directivity at wavelengths less than the total length of the array *l*, but in order to have good control over both beamwidth and direction the following condition is required:

$$l \gg \lambda \tag{3.51}$$

In fact, an array five times longer than the wavelength is enough to meet the requirement [3]. The minimum frequency is calculated as follows.

$$l = N \cdot d = 0.975\,m$$

$$l \gg \lambda \quad \rightarrow \quad l = 5 \cdot \lambda \rightarrow \lambda = 0.195\,m$$

$$f_{\min} = \frac{v_s}{\lambda} = 1743.6\,Hz \tag{3.52}$$

The following figures show the Array Factor (AF) of the signal transmitted by the sound projector depending on the frequency. Figure 3.4 shows the beam pattern of two sinusoids with a central frequency of 1743 Hz and 5056 Hz respectively. As explained above, these frequencies are the limits of the available frequency range for the loudspeaker array.

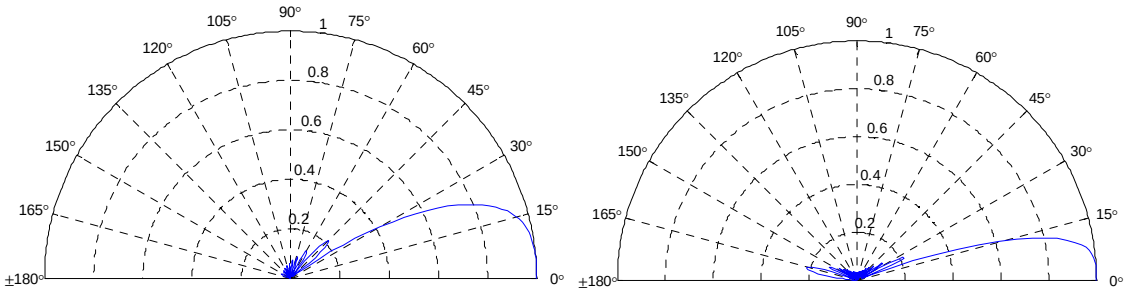



**Figure 3.4.** Array Factor for ULA ($N$ = 30, $d$ = 0.0325 *m*, $\theta$ = 0º) using frequencies of 1743 *Hz* and 5056 *Hz* respectively.

Figure 3.5 shows the beam pattern of two sinusoids with a central frequency of 500 Hz and 6500 Hz respectively. In the first graph it is possible to see how the main lobe increases its

beamwidth dramatically. In the second, a grating lobe, steered in the direction of 128º approximately, affects the beam pattern.

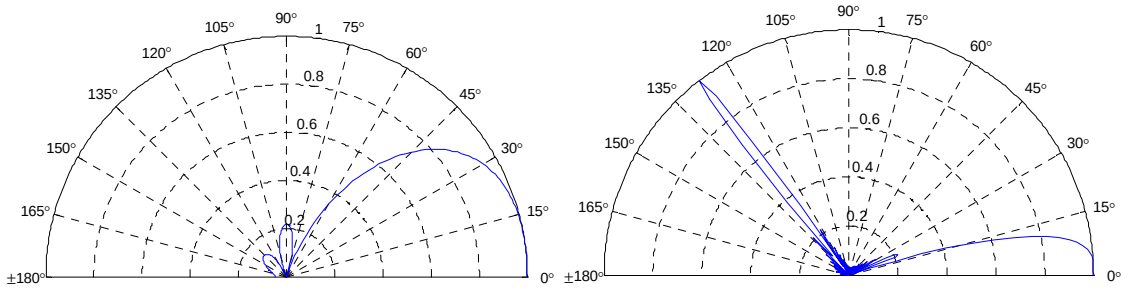



**Figure 3.5.** Array Factor for ULA ($N = 30$, $d = 0.0325\ m$, $\theta = 0º$) using frequencies of 500 *Hz* and 6500 *Hz* respectively.

### 3.5.2 Frequency Range Limited by the Matrix of Microphones

The configuration of the sensors only affects the maximum frequency of the signal they can receive in order not to encounter aliasing.

In this case, the microphone constellation is composed by 4 microphones set up in a square. This matrix can be decomposed in six pairs of microphones. The minimum distance between the sensors limits the frequency.

In order to find the maximum frequency, the analysis done for the loudspeaker array in the previous section can be followed here as well. In this case, the array consists of two elements, so the condition (3.49) restricts the available range more strongly.

Comparing both constraints, (3.46) and (3.49), the respective frequencies can be obtained:

$$d < \frac{\lambda}{2} \quad \rightarrow \quad f_{\max} = \frac{v_s}{\lambda} = \frac{v_s}{2 \cdot d} = 4250\,Hz \tag{3.53}$$

$$\lambda > \frac{2d \cdot N}{N-1} \quad \rightarrow \quad f'_{\max} = \frac{v_s}{2 \cdot d} \cdot \left( \frac{N-1}{N} \right) = 2125\,Hz \tag{3.54}$$

As it can be observed, the available range of frequencies varies noticeably from choosing (3.53) or (3.54). Therefore, the maximum frequency cannot be determined theoretically and it has to be found experimentally.

# Chapter 4

# Signal Description

Because the sounding and projector phase have different purposes, the features of the transmitted signal must be also different in each phase. The sounding phase needs the design of a specific and particular signal; meanwhile in the projector phase, the signal is obtained from the DVD-player and only the bandwidth can be modified.

In this section an analysis is developed in order to find the properties of both signals.

## 4.1 Design of the Signal Used during the Sounding Phase

### 4.1.1 Wideband Additive Gaussian Noise Signal

The selected signal to perform the tests during the sounding phase is a wideband additive Gaussian noise signal.

The decision was taken by knowing that the implemented methods need uncorrelated signals to work properly (see section 3.3). They are based on the cross-correlation between the received signals, so the received signals need to be as uncorrelated as possible while meeting the requirements seen in subsection 3.5.

### 4.1.2 Length in Samples

Due to the length of the transmitted and received signals must be the same, the minimum length of the signal is limited by the maximum size of the test room. The choice of this number fulfilled two constraints:

- In case of running the sound projector in a room the signal should be able to arrive to the matrix of microphones before its transmission ends. So, to cover a maximum of 15 meters squared room a long sampled signal is needed, with a minimum of:

$$f_s = 48000\,Hz, \quad d_{\max covered} \simeq 25\,m$$

$$\left\lceil \frac{d_{\max covered} \cdot 2}{v_s} \cdot f_s \right\rceil = 7059\,samples \tag{4.1}$$

- The digital interface between the A/D converters and the DSP set the input and output buffers to 60 samples (see subsection 5.2.1). So, in order to transmit the noise signal, a total length multiple of 60 was selected.

Finally, a length of 8220 samples was chosen.

### 4.1.3 Selected Bandwidth

After developing the theoretical analysis in subsection 3.5, several experimental tests were done in order to find a good tradeoff between not losing too much bandwidth and not having aliasing effects at the reception.

For the signal used during the sounding phase, the minimum frequency is given by the loudspeaker array; meanwhile, the maximum is given by the sensors.

Finally, the selected bandwidth included from $f_{c_1} = 1700\,Hz$ to $f_{c_2} = 3500\,Hz$, based on the analysis in subsection 3.5.

### 4.1.4 Final Design

The selected signal is a band-pass additive Gaussian noise. The signal used during the sounding phase is shown in Figure 4.1. Its parameters are:

- Band-pass-filtered additive Gaussian noise.
- Bandwidth =1800 Hz ( $f_{c_1} = 1700\,Hz$ , $f_{c_2} = 3500\,Hz$ ).
- Central frequency = 2600 Hz.
- Length: 8220 samples (which is equal to 171.25 ms of duration).

The signal was created and filtered using the `randn()`and `filter()` functions in MATLAB and, later, loaded into the DSP.

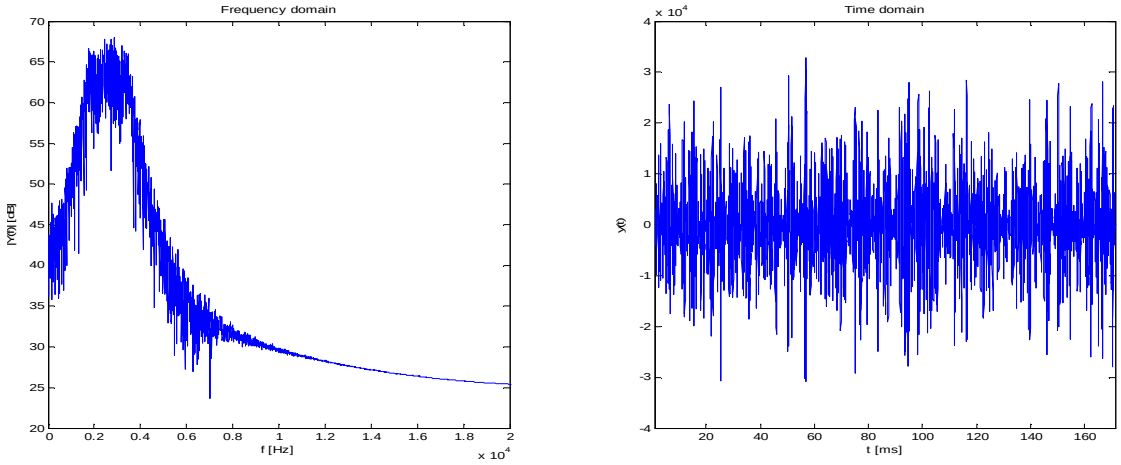



**Figure 4.1.** Frequency and time domain of the transmitted signal.

## 4.2 Modifications of the Signal Obtained During the Projector Phase

The only modification applied to the signal obtained from the DVD-player is a filtering in order to avoid the higher frequencies.

### 4.2.1 Selected Range of Frequencies

For the projector phase, the limitations obtained from the microphones have no sense due to they are not connected to the SAM-30 and a human listener is located instead of them. So, only the loudspeaker array limitations will affect in this phase (see subsection 3.5).

After several acoustic experiments with human listeners, the selected bandwidth was from 0 *Hz* to 5000 *Hz*. It was observed that the frequency range from 0 *Hz* to 1700 *Hz* modified very little the perception of the direction of arrival of the sound; while for frequencies higher than 5000 *Hz* a certain ambiguity was observed in the definition of the perception of the mentioned direction.

# Chapter 5

# Implementation

The implementation of the system is divided in several modules. These modules correspond to the three phases mentioned above, namely sounding, projector and graphics phase. The programs which implement the sounding and projector phases run on the DSP and are written in C; while the program which implements the graphics phase is written in MATLAB. The graphical user interface (GUI), which provides the communication between the PC and the DSP, is also written in MATLAB.

## 5.1 Necessary Parameters to Run the Projector Phase

Three parameters (`steering_delay`, `attenuation` and `transmit_delay`) control the projector phase. They are calculated during the sounding phase in the SAM-30 or during the post-analyses in MATLAB. Their meaning is the following:

- **`steering_delay`:** It is an array with four positions. It contains the delay to be applied to the transmitted signals between consecutive speakers in the projector phase. Each direction of transmission corresponds to one channel in a 4.0 setup (front right/left, and rear right/left channels). In the Beamforming subsection (see subsection 3.1.1), this parameter is called $D_l$ ($l$=1…4).
- **`attenuation`:** It is an array with four positions. It contains the attenuation to be applied to the four channels during the projector phase. This variable allows the listener to hear the four channels with the same volume. It is calculated from the sound power received from each channel. In the Beamforming subsection, this parameter is called $a_l$ ($l$=1…4).

- **`transmit_delay:`** It is an array with four positions. It is used during the projector phase in order to synchronize the signals aimed for the four channels at the position of the listener. It indicates the number of samples which each channel has to delay its signal. It is used in order to suppress the propagation delay due to the path followed for the signal in each channel. In the Beamforming subsection, this parameter is called $T_l$ ($l$=1…4).

## 5.2 DSP

Code Composer Studio v.3.3 (CCS) is the software used to implement the algorithms and load code onto a DSP. CCS provides integrated program management using projects. A project keeps track of all information that is needed to build a target program.

Thus, the sounding and projector phase are implemented in separated projects (Sounding_phase_48kHz.pjt and Projector_phase_48kHz) in the DSP. In the following sections their features and structure are explained.

Both Sounding_phase_48kHz.pjt and Projector_phase_48kHz.pjt were implemented based on a previous program (loudspeaker_array_48kHz.pjt) whose code was reused. It contains a basic skeleton of the program which initializes the DSP and manages the input and output buffers in real time. Some of the files which make up the structure of this program are explained below:

- **`application.c`**: This file is composed of three empty functions (`appl_init()`, `appl_process()` and `appl_Background_process()`) which have to be filled with the code that implements the system's functionality.
- **`edma_mcbsp_quad_loudspeaker_setup.c`**: Sets up the I/O buffering.
- **`loudspeaker_array_setup.c`**: Sets up the specific functionality for the 30-element loudspeaker array.
- **`main.c`**: This file contains the `main()` function. It initializes the DSP and the application, and leaves the program running in a loop which calls `appl_process()` each time that new samples are received in the I/O buffers.

This basic skeleton is prepared in order not to be modified; it provides a structure that maintains the program in an infinite loop preparing the input and output buffers to be manipulated. In consequence, only `appl_process()`, `appl_background_process()` and `appl_init()`have to be filled in order to implement the functionality of the system.

The libraries used in the DSP code are the ones that come with Code Composer Studio and the standard C library, listed below.

- csl6713.lib (chip support library for the C6713 processor).
- dsp67x.lib (collection of 64 high-level optimized DSP functions for the TMS320C67x device).
- rts6700.lib (run-time support for the C67x architecture).

### 5.2.1 I/O Buffers

The input and output buffers are prepared to be used easily by the functions included in the `edma_mcbsp_quad_loudspeaker_setup.c` and `loudspeaker_array_setup.c` files. As explained above, these files were already written and its code was reused for this application.

Direct memory access (DMA) is used to store the data into a temporary input buffer. An interruption is passed to the processor when the input buffer becomes full. Upon receiving the input buffer the `application_process()` function fills the output buffer with the samples which will be sent through the loudspeaker array. If this is done before the next input buffer arrives, the program will run in real time.

The input and output buffers are shown in Figure 5.1 and Figure 5.2. They differ in their length. These buffers are declared as `short` type and their lengths are 240 and 1800, respectively. The input buffer receives 60 samples from each microphone, being four in total. The output buffer sends 60 samples to each loudspeaker, being thirty in total.

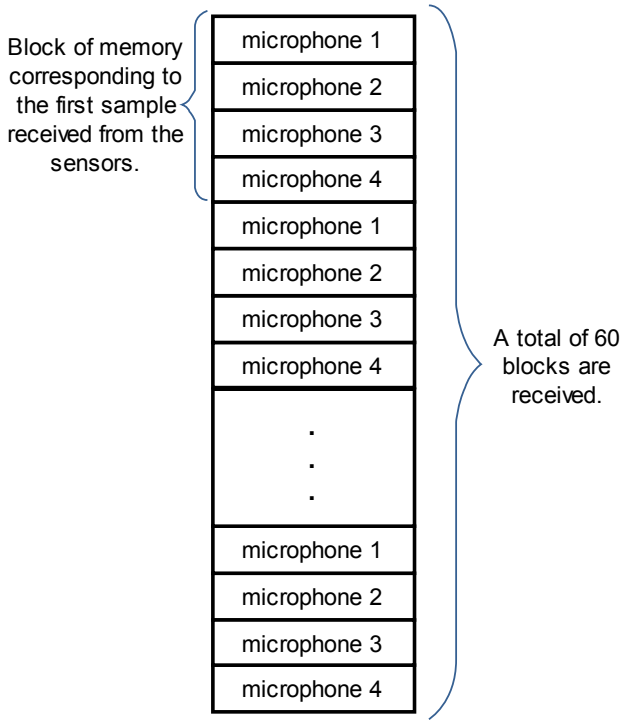



**Figure 5.1.** Representation of the input buffer.

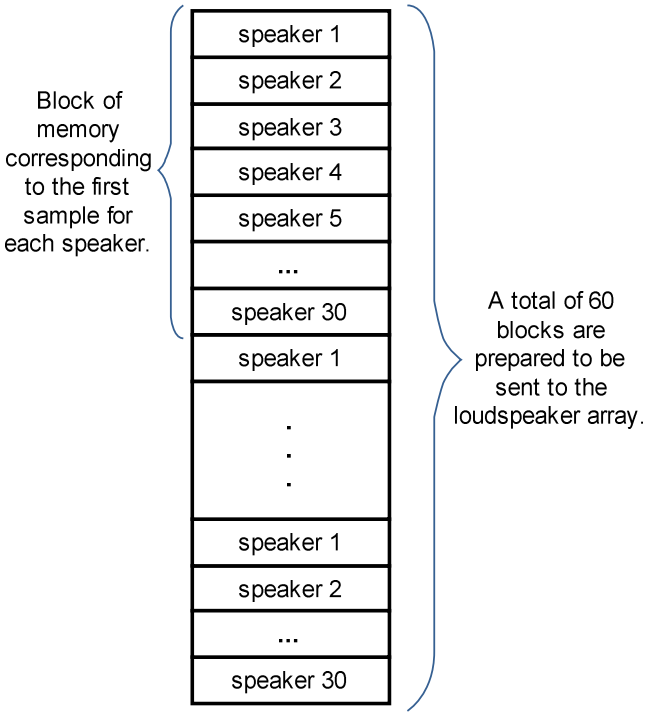



**Figure 5.2.** Representation of the output buffer.

### 5.2.2 Sounding_phase_48kHz.pjt

The Sounding_phase_48kHz.pjt is the program run during the sounding phase. The aim of this program is to produce the `steering_delay`, `attenuation` and `transmit_delay` parameters (see subsection 5.1) of each channel for a specific room. The TDOA method used in this program is the CC technique explained in subsection 3.3.2.1 (Search for the TDOA between Signals). The CC technique does not support the calculation of the `transmit_delay` variable, so their values are set to zero in MATLAB (in the laptop PC).

In order to obtain the `steering_delay` and `attenuation` variables, the signal described in section 4.1 is transmitted. The loudspeaker array does a sweep, sending the signal in a number of fixed directions. While the sound projector transmits it, the microphones are already recording. Both transmission and reception of the samples is done in real time. Afterwards, for each direction of transmission, an analysis is done in order to estimate the DOA, the sound power received from this DOA and the sound power received from an angle of 0º (direct sound). Then, these values are stored in the sounding table.

The sounding table consists of four columns and nineteen rows. The nineteen rows correspond to the directions of transmission. The columns are composed by the following values:

- Direction of transmission.
- Direction of arrival (DOA).
- Sound power from the DOA.
- Sound power from an angle of 0º.

| | **Direction of transmission $\varphi$[*degrees*]** | **DOA $\theta$[*degrees*]** | **Sound power from the DOA $P(\theta)$** | **Sound power from 0º $P(0º)$** |
|---|---|---|---|---|
| **1** | $11.3^{o}$ | $349^{o}$ | $8.4 \cdot 10^{5}$ | $2.4 \cdot 10^{4}$ |
| **2** | $29.3^{o}$ | $330^{o}$ | $1.1 \cdot 10^{5}$ | $4.5 \cdot 10^{4}$ |
| **…** | … | … | … | … |
| **19** | $168.7^{o}$ | $20^{o}$ | $7.8 \cdot 10^{5}$ | $3.2 \cdot 10^{4}$ |

**Table 5.1. Example of the possible values in the sounding table. The values are not real.**

Once all the angles and powers are stored in the sounding table, the best ones with relation to each channel (front left/right, and rear left/right) are kept. Finally, the `attenuation` and `steering_delay` variables are obtained and returned to MATLAB (in the laptop PC).

Sounding_phase_48kHz.pjt adds some new files in the structure. The most important functions and features are explained below:

- **`application.c`:** This file already existed, but it was totally rewritten. It contains the `appl_init()` function, which is called at the beginning of the program to perform all

the initialization, as well as `init_next_step()` and `init_next_xcorr()`, which initialize some variables in order to send the signal in the next direction or to calculate the next cross-correlation. There is also `appl_process()`, which controls the process of the application. This is done using a switch which determines the current state and runs the necessary functions in order to do it.

- **`correlation.c`:** All the functions included in this file are related to cross-correlation and power terms. The `xcorr_init()` function initializes the global variables used in this file. Functions like `calculate_power()` and `calculate_power_0()`, as their names suggest, are devoted to obtaining the sound power from a specific direction. Moreover, `xcorr2()` calculates the cross-correlation between two signals, and `do_xcorr2()` manages the cross-correlation process.
- **`directions.c`:** The functions which compose this file are related to the DOA. `get_TDOA()` gives the TDOA between signals recorded on a pair of microphones. `get_TDOA_diag()` gives the TDOA related to the two pairs of microphones which form the diagonals converted to the Cartesian plane. Finally, the `get_angle()` function gives the angle related to the DOA.
- **`equalizer.c`:** It contains the necessary functions to choose the `attenuation` and `steering_delay` variables. Function `equalizer_init()` does the initialization. Function `calculate_eq_coef()` obtains the `attenuation` and `steering_delay` variables, and `equalize()` manages the equalization process.
- **`interpolation.c`:** This file has only two functions: `interpolation_init()`, which initializes the global variables used in this file, and `interpolate2()`, which interpolates the incoming buffer by 2. This function is used for the beamforming.
- **`micr_calibration.c`:** It contains the functions needed to calibrate the matrix of microphones. Function `micr_cal_init()` initializes the global variables used in this file, `calculate_cal_power()` and `calculate_calibration_coef()` are used to obtain the power and the calibration coefficients, respectively, and `do_microphone_calibration()` manages the calibration process.
- **`signal_IO.c`:** This file contains functions related to the beamforming and the storage of the input and output samples. The `prepare_loudspeaker_buffer()` function, as its name suggests, prepares the signal to be loaded in the loudspeaker buffer, since these signals have to be beamformed digitally before being transmitted through the sound projector. Function `send_no_signal()` is used while the analysis is running and no sound is wanted through the speakers. Dealing with the storage of the samples, `store_buffer()` keeps the current input buffer in a specific range of memory on the DSP.

Figure 5.3 illustrates the flowchart of this program. When it starts, the DSP and all the variables are initialized. The variable which controls the process of the application is called `state` and it is initialized to `SEND_CALL_SIGNAL`, which is the first state to perform. Then it goes into a loop which finalizes when this variable achieves the `END` value. In order to control the process, a switch structure is needed. Depending on in which state the program is, different functions have to be called from `appl_process()`.

The first state the application enters is `SEND_CALL_SIGNAL`. In this state, the signal is sent directly from the loudspeaker array towards the microphone constellation, without doing any beamforming.

Once the signal is transmitted and also received, the `state` variable changes to `CALIBRATION`. Here, with `do_microphone_calibration()`the received power of the signals in each microphone is calculated and the calibration coefficients are obtained. This calibration is needed in order to make the received power in each sensor equal, since the batteries of the microphones might not be at the same level.

At this point, the `state` variable acquires the `RUNNING` value and the program is ready to begin the sounding phase. The signal is transmitted through the loudspeaker array and steered towards a specific direction (see Table 3.1). The beamforming of the signal is done by using the `prepare_loudspeaker_buffer()` function. Moreover, the samples recorded by the microphones are stored in a specific range of memory using `store_buffer()`. They are kept in the memory of the DSP in order to be loaded on the PC for subsequent analysis in MATLAB.

After the transmission and reception of the signal, the analysis begins. The `state` variable achieves the `XCORR` value. This analysis performs the CC technique explained in section 3.3.2.1. The first part of it consists in cross-correlating the four recorded signals in pairs in order to find the position of the maximum. This is done by `do_xcorr2()`. A total of six cross-correlations have to be done in this state and for this purpose `xcorr2()` and `DSPF_sp_dotp_sqr()` are used.

Once the six cross-correlations are done, the DOA is calculated. Here, `state` gets the `ANGLES` value. The position of the maximum in each one of these cross-correlations gives the TDOA between every two signals recorded on the microphones. These delays are obtained with `get_TDOA()` and `get_TDOA_diag()`. The TDOAs obtained from adjacent pairs of microphones are already given in Cartesian plane, but those obtained from the diagonals must be converted. So, `get_TDOA_diag()` gives the delay of the diagonals converted to this reference system. Knowing those values, the angle of arrival is obtained with `get_angle()` (as seen theoretically in subsection 3.4.1).

At this moment, the angle which gives the maximum sound power for the current direction of transmission is already calculated, but the analysis still requires the knowledge of other power values. Following the progress of this program, now the value of `state` is `POWER_0`. Here, the sound power received from 0º (direct sound) is obtained using `calculate_power_0()`.

In the same way, the sound power received from the DOA is needed. It is obtained when the `state` variable is `POWER`. The `calculate_power()` function performs this calculation using the values obtained from the previous cross-correlations. These values (angle of arrival, sound power from the DOA and sound power from 0º) are stored in a table with its corresponding direction of transmission.

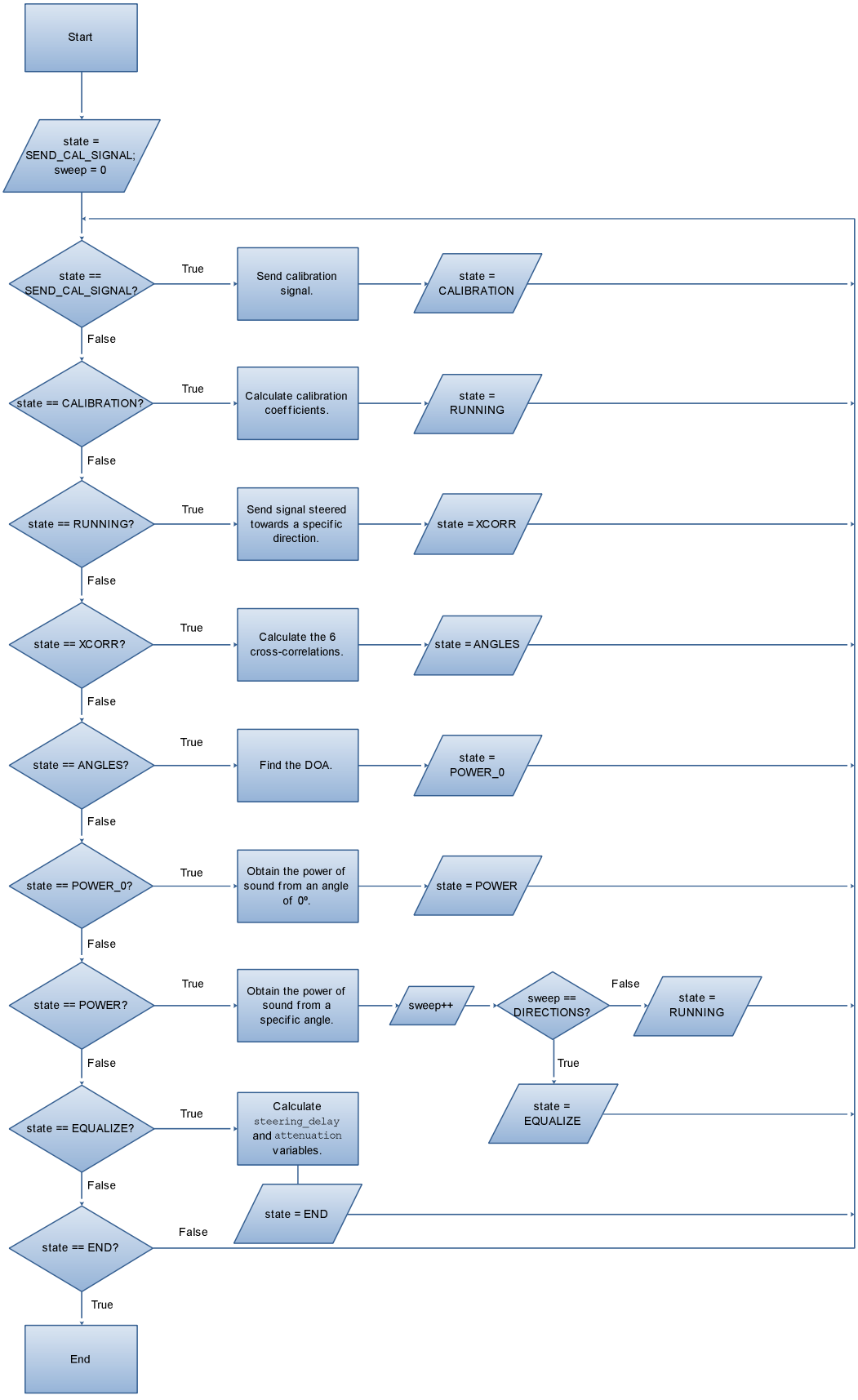



**Figure 5.3.** Flowchart of Sounding_phase_48kHz.pjt.

When the analysis achieves this point, the `sweep` variable is incremented and the signal is ready to be sent towards the next direction of transmission. Now, `state` acquires the `RUNNING` value and the analysis is repeated for the next direction of transmission. When the sweep is completed, or in other words when `sweep` gets the `DIRECTIONS` value, the sounding table is already filled.

Thus, `state` achieves the `EQUALIZE` value and the `equalize()` function is called. The values of each row of the sounding table are used in this function leading to 19 new values. These values are the difference of sound power between the DOA and the direction of 0º and they are calculated as

$$L = 10 \cdot \log\left(\frac{P(\theta)}{P(0^{\circ})}\right) \tag{5.1}$$

where $P(0^{\circ})$ is the sound power received from 0º and $P(\theta)$ is the sound power received from the DOA.

Then, the angles of arrival are classified according to the quadrant they belong in the coordinate system of the matrix of microphones. Each quadrant is related to one channel (see Figure 2.1):

- Channel 1: Front right. From 270º to 360º. The preferred DOA is 315º.
- Channel 2: Rear right. From 180º to 270º. The preferred DOA is 225º.
- Channel 3: Rear left. From 90º to 180º. The preferred DOA is 135º.
- Channel 4: Front left. From 0º to 90º. The preferred DOA is 45º.

For each channel, the difference of sound power (see equation (5.1)) which gives the highest score is chosen. This value is selected along with all the values in the row of the sounding table (direction of transmission, angle of arrival, sound power from the DOA and sound power from 0º). Thus, from this selection, the `steering_delay` and `attenuation` variables are calculated. `steering_delay` is obtained from the direction of transmission (see Table 3.1), and `attenuation` is obtained as

$$P_{\min} = \min\left(P_1(\theta), P_2(\theta), P_3(\theta), P_4(\theta)\right) \tag{5.2}$$

$$a_i = \sqrt{\frac{P_{\min}}{P_i(\theta)}} \qquad i = 1...4 \tag{5.3}$$

where $P_i(\theta)$ ($i$=1…4) are the selected sound powers received from the DOA for the channel $i$, and $a_i$ are the attenuation to be applied to the output of the loudspeaker array during the projector phase.

Finally, the `END` state is reached and the program finalizes returning the `steering_delay` and `attenuation` variables to MATLAB (in the laptop PC).

### 5.2.3 Projector_phase_48kHz.pjt

After running the sounding phase (with the Sounding_phase_48kHz.pjt program in the SAM-30), the analysis of the room is already done and the variables (`steering_delay`, `attenuation` and `transmit_delay`) which indicate the optimum parameters for the transmission of each channel are obtained. Thus, the projector phase can begin: the Projector_phase_48kHz.pjt program can be loaded and run on the SAM-30.

The structure of this program follows the structure of the previous one, but it is much simpler. Figure 5.4 illustrates its flowchart. Only two states are needed: `RUNNING` and `END`.

Projector_phase_48kHz.pjt reuses the some functions and files created in Sounding_phase_48kHz.pjt. The most important functions and features are explained below:

- **`application.c`:** This file already existed, but it was totally rewritten. It contains the `appl_init()` function, which is called at the beginning of the program to perform all the initialization. There is also `appl_process()`, which controls the process of the application.
- **`interpolation.c`:** This file has only two functions: `interpolation_init()`, which initializes the global variables used in this file, and `interpolate2()`, which interpolates the incoming buffer by 2. This function is used for the beamforming.
- **`signal_IO.c`:** This file contains functions related to the beamforming. The `prepare_loudspeaker_buffer()` function, as its name suggests, prepares the signal to be loaded in the loudspeaker buffer, since these signals have to be beamformed digitally before being transmitted through the sound projector.

First, the DSP and all the parameters are initialized. Then, the program enters a loop which finalizes when `state` takes the `END` value. Since there are only two states, the `state` variable is initialized as `RUNNING`. This loop calls `appl_process()` each time new samples are received from the DVD-player.

The first step is to filter the four channels in order to avoid frequencies higher than 5050 *Hz* (due to the limitation shown in section 3.5.1). Otherwise, the frequencies higher than this value will be steered towards other directions of transmission and this could affect the sound effect expected at the position of the listener. This is done with the `FIR_filter()` function. Thus, the chosen solution was focused on implementing a $19^{th}$-order low-pass FIR filter with a cut-off frequency of 5050 *Hz*.

Then, `prepare_loudspeaker_buffer()` is called and the samples are beamformed using the method described in subsection 3.2 and sent to the sound projector. Moreover, in order to run the application in real time, the samples must be prepared to be transmitted through the loudspeaker array before receiving the new samples from the DVD-player.

The program repeats this procedure until the `END` state is set by the user.

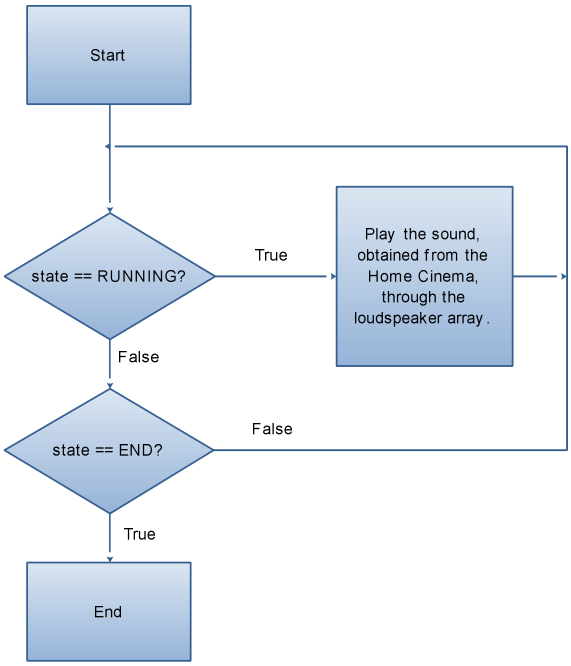



**Figure 5.4.** Flowchart of Projector_phase_48kHz.pjt.

## 5.3 MATLAB Interface and Post-Analysis

Due to the features of MATLAB and the inclusion of the Link for CCS Development Tools, the graphical user interface (GUI) and post-analysis were developed in MATLAB.

### 5.3.1 GUI: "Loudspeaker Array Manager"

The developed GUI controls the whole system. It lets the user begin the sounding phase, select another kind of signal to transmit, analyze the results in depth by showing graphics and figures and, finally, run the projector phase in three possible modes: mono, stereo or surround. It also allows the user to load the signals stored from previous sessions and analyze the results graphically.

The Loudspeaker Array Manager initializes Code Composer Studio by itself and loads the necessary files to run the program, so the user does not need to run any other program apart from MATLAB.

In Appendix B, there is a tutorial explaining how to use the "Loudspeaker Array Manager".

### 5.3.2 Graphics Phase

In order to perform these post-analyses in MATLAB, the recorded samples must be loaded to MATLAB from the SAM-30. All the options are shown in Appendix B, in this subsection only the options related with both TDOA techniques are commented.

The graphics shown in this phase are:

- Transmitted and received signals in time and frequency domain.
- Graphics corresponding to CC and CSS methods.
- Beam pattern of the loudspeaker array for each specific direction of transmission.

Since these post-analyses do not have to meet the real-time condition, its implementation lacks interest and only some specific points are explained.

The graphical presents some values and results that are not needed for the implementation of the system. Bearing in mind what has been explained in chapters 3.3 and 3.4, in the following subsections, it is explained how the DOA plots are obtained in MATLAB.

#### 5.3.2.1 Post-Analysis Performed Using the CC Technique

The CC technique is already performed in the SAM-30 during the sounding phase. The variables obtained in MATLAB (see subsection 5.1) have the same values that the ones calculated in the SAM-30. It also sets the `transmit_delay` variable to zero.

However, in this case and thanks to MATLAB, it is possible to show some graphics:

- Radiation pattern received at the microphone constellation for a specific direction of transmission. This polar plot was obtained in the way the following pseudo code shows. This technique has been theoretically explained in subsection 3.3.2.2.

```
angles = 0:360/(n_points):360;                  %Define the DOAs.

for i = angles                                  %For each DOA.
   delays = get_delays(angle(i));               %Obtain delays from the specific DOA.
   m_t = apply_delays(m1, m2, m3, m4, delays);  %Apply these delays to the recorded
                                                % signals and add them.
   power(i) = get_power_from_delays(m_t);       %Obtain the power in this DOA.
end

polar(angles, power);                           %Show the polar graphic.
```

**Figure 5.5.** Pseudo-code corresponding to the polar plot obtained from the CC technique.

- Time-DOA plot of the received instantaneous sound power during the sounding phase. Next is a plot which shows the received instantaneous sound power according to the

DOA as a function of delay. It is a 3D plot. This technique has been theoretically explained in subsection 3.3.2.3.

```
time = convert(samples);                          %Convert the samples to time.

for i = angles                                    %For each DOA.
   delays = get_delays(angle(i));                 %Obtain delays from the specific DOA.
   m_t = apply_delays(m1, m2, m3, m4, delays);    %Apply these delays to the recorded
                                                  % signals and add them.
   power(i) = get_instantaneous_power(m_t,s,);    %Get the instantaneous power for each
end                                               % sample.

mesh(angles, time, power);                        %Show a 2D graphic depending on the
                                                  % DOA and the time.
```

**Figure 5.6.** Pseudo-code corresponding to the time-DOA plot of the received instantaneous sound power obtained from the CC technique.

#### 5.3.2.2 Post-Analysis Performed Using the CSS Technique

The CSS technique is only implemented in MATLAB. In this case, it must be emphasized that, due to the cross-correlation between the transmitted and the four received signals, this technique allows the calculation of all three variables (see subsection 5.1): `steering_delay`, `attenuation` and `transmit_delay`.

This method obtains two different figures:

- Propagation delay-DOA plot for a specific direction of transmission. It is a 3D plot which shows the value of the cross-correlation between the transmitted and the received signals in function of the DOA and the propagation delay. This technique has been theoretically explained in subsection 3.3.3.1.

```
angles = 0:360/(n_points):360;                    %Define the DOAs.
time = convert(samples);                          %Convert the samples to time.

for i = angles                                    %For each DOA.
   delays = get_delays(angle(i));                 %Obtain delays from the specific DOA.
   m_t = apply_delays(m1, m2, m3, m4, delays);    %Apply these delays to the recorded
                                                  % signals and add them up.
   xcorr(i) = get_xcorr_from_delays(m_t,s);       %Get the cross-correlation between the
end                                               % recorded and transmitted signals.

mesh(angles, time, xcorr);                        %Show a 3D plot depending on the
                                                  % DOA and the propagation delay.
```

**Figure 5.7.** Pseudo-code corresponding to the propagation delay-DOA plot obtained from the CSS technique.

- Radiation pattern received at the microphone constellation for a specific direction of transmission. It sums up the information given in the previous graphic on a polar plot, losing the time information. It selects the values higher than the background noise of the previously obtained cross-correlation and shows them in a polar plot. This technique has been theoretically explained in subsection 3.3.3.2.

```
level = calculate_otsu_threeshold(xcorr)        %Computes a global threshold using
                                                % Otsu's method.

for i = angles                                  %For each DOA.
   for j = samples                              %For each sample.
      if( xcorr(i,j)>level )                    %Compare the current value of the
         pattern(i) = pattern(i)+xcorr(i,j);    % cross-correlation done in Figure 5.7
      end                                       % if this value is higher than level,
   end                                          % add it to the pattern.
end

polar(angles, pattern);                         %Show the pattern in a polar graphic.
```

**Figure 5.8.** Pseudo-code corresponding to the polar plot obtained from the CSS technique.

# Chapter 6

# Results

A number of simulations have been carried out in order to experimentally study the two algorithms (CC and CSS techniques) in real acoustic environments under different operation conditions.

In this section, a brief account of them is presented. They highlight the merits and limitations inherent in these TDOA techniques and justify what it has been learned through theoretical analysis in the previous sections.

It is structured in two main subsections: the objective results, obtained from the samples recorded on the microphones, and the subjective impressions, which are the feelings given from the sound projector to human listeners. An explanation of the meaning of some plots is given.

All the simulations were conducted in the laboratory room described in subsection 2.1. Keeping in mind the position of the loudspeaker array in this room, the signals obtained from directions of transmission less than 90º (see Figure 2.1) were very poor and they lost almost all their power so that their DOA cannot be found. Meanwhile the signals obtained from directions of transmission greater than 90º gave good results and show their DOA properly.

This way, during the analysis of the results, the graphics and plots obtained from directions of transmission less than 90º will not be shown.

Additional graphs and plots appear in Appendix D.

## 6.1 Objective Results

This section gives an analysis of the results obtained from the graphics phase.

### 6.1.1 Tests

Several different tests were done to analyze the behavior, the robustness and the effectiveness of the TDOA methods explained in subsection 3.3.

They can be divided in two groups:

- The first group included all the tests performed to find the best features of the transmitted signal for the described hardware in order to obtain good results with the proposed TDOA techniques. These tests helped to find the selected parameters which characterize the transmitted signal (see section Chapter 4).
- Once the signal parameters were decided, the second group of tests focused on running the sounding phase of the system. Two situations were taken into account:
    - A quiet environment. In this case, the optimal conditions are considered. It emulated the typical set-up situation in which no extern noise or interfering signals disturb it.
    - A noisy environment. This situation focused on showing the robustness of both TDOA techniques under the effect of interfering signals. For this purpose, another speaker was located at 80º from the matrix of microphones. The emitted interfering signals were an additive Gaussian noise with different amplitude levels. Then, the sounding phase was run while the new speaker transmitted one of those signals. This procedure was repeated for all the interfering signals.

The most relevant results were obtained from the second group of tests. In the following subsections, some of these results are analyzed:

- Ten of the total tests conducted in the quiet environment. These tests gave very similar results and behavior. In the following subsections, all the graphics shown correspond to one of the ten tests arbitrarily selected.
- Five of the total tests conducted in the noisy environment. The interfering source was located at 80º and 4.3 *m* away from the matrix of microphones. The interfering signal was an additive Gaussian noise and it was 1.3 *dB* lower in amplitude than the true one (the signal transmitted by the loudspeaker array). As it happens in the previous tests, these ones also gave very similar behavior, so that all the graphics only correspond to one of them.

These tests are analyzed in function of the TDOA techniques.

### 6.1.2 Results Obtained from the CC Technique

As it is explained in section 3.3.1, the CC technique only takes into account the samples recorded on the microphones and knowledge of the sent signal is not used for the calculations. The following subsections include the results of the tests done in both environments and the analysis and discussion of their behavior for this TDOA technique.

#### 6.1.2.1 Tests Performed in a Quiet Environment

After running the ten tests, the samples were already loaded to MATLAB in order to graphically verify the values given by the SAM-30. For each test, the SAM-30 returned two variables (`steering_delay` and `attenuation`) which are used during the projector phase. The `transmit_delay` variable is set to zero for this technique.

These two variables are described in subsection 5.1, and they correspond to the attenuation and delays necessaries at the loudspeaker array to send the sound waves of each channel (front left/right and rear left/right) to the correct direction. See Table 3.1 in order to observe the relation between the delay applied at the sound projector and the direction of transmission.

Due to the left wall was located at 9 *m* from the loudspeaker array (see subsection 2.1), the sound power obtained from the sound waves steered towards that wall (i. e. when the direction of transmission is less than 90º) is very low. Therefore, the graphics and values obtained from these directions of transmission do not show any clear result. Thus, the values of `steering_delay` and `attenuation` given for channels 3 and 4 are useless. All the values obtained from the ten experiments are detailed in Appendix C.

The average values of `steering_delay` and `attenuation` for channels 1 and 2 are shown in the following table.

| Variable | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **`steering_delay`** | -3.5 | -2 |
| **`attenuation`** | 0.6560 | 1.0000 |

**Table 6.1.** Average values of `steering_delay` and `attenuation` obtained from the CC technique for channels 1 and 2 in a quiet environment.

Bearing in mind Figure 2.1, the meaning of the values of the `steering_delay` is that for channel 1, which corresponds to an angle of arrival at the sensors between 270º and 360º, the delay at the beamformer that delivers the best quality is -3.5 samples; in the same manner, for channel 2, which corresponds to a direction of arrival at the sensors between 180º and 270º, the best delay is -2 samples. According to Table 3.1, a steering delay of -3.5 samples corresponds to an angle of transmission of 139.7º at the beamformer. Similarly, a steering delay of -2 samples corresponds to 115.8º.

The `attenuation` variable is used to make the power equal for all channels at the position of the listener. In order to do it, the most powerful channels have to be attenuated. From these values, channel 2 does not apply any attenuation and the amplitude of channel 1 must be attenuated to an average of 66%.

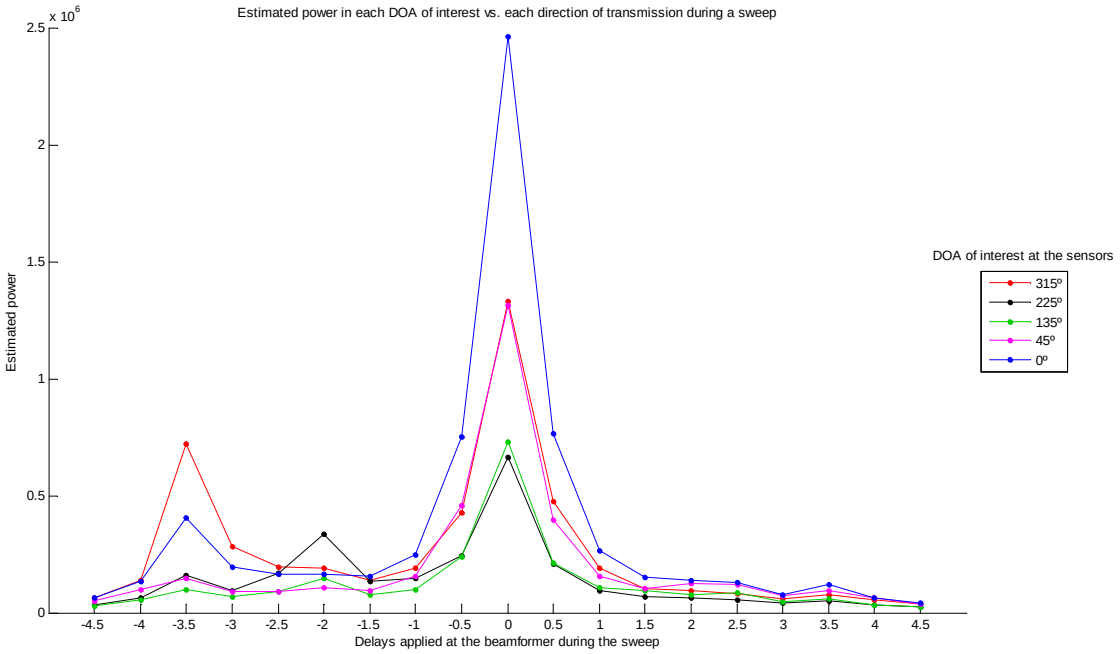



**Figure 6.1.** Example of the estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) in a quiet environment. It has been obtained from the CC technique.

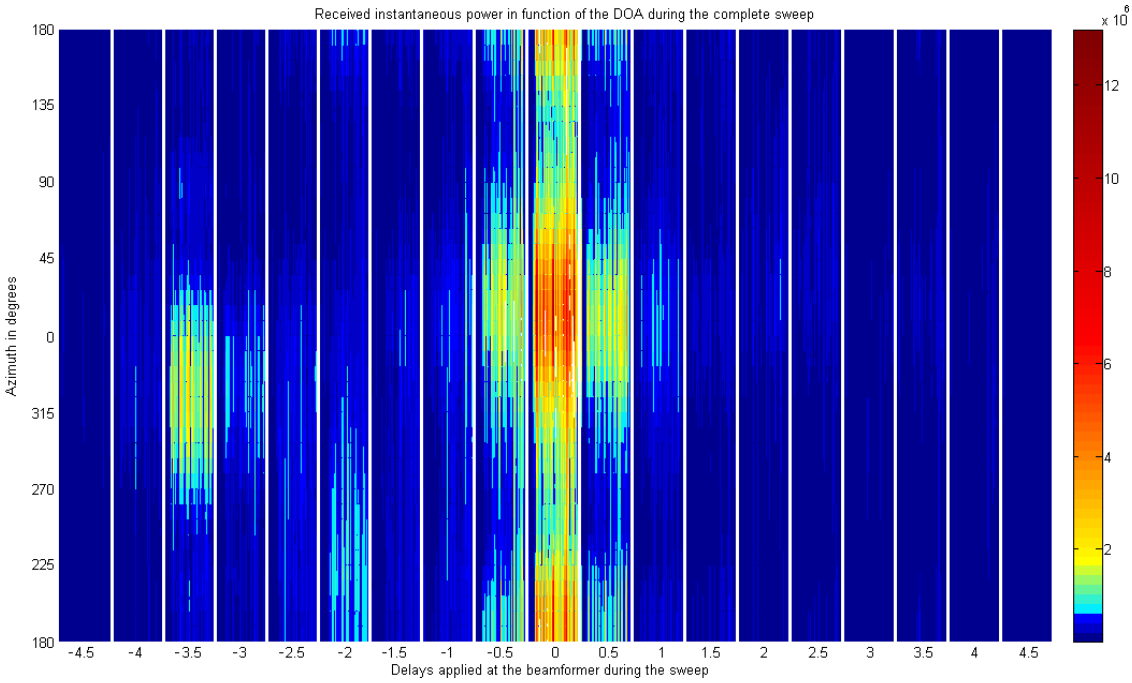



**Figure 6.2.** Time-DOA plot of the received instantaneous sound power during the sounding phase in a quiet environment. It has been obtained from the CC technique.

Figure 6.1 shows the sound power received in the directions of interest during the sounding phase (red, black, green, pink and blue correspond to 315º, 225º, 135º, 45º and 0º respectively). The graphic has been arbitrarily given from one of the ten tests performed under this environment. The vertical axis refers to the sound power and the horizontal axis refers to the delay applied at the loudspeaker array in order to transmit the signal (see Table 3.1). During the sounding phase, the sound projector steers the main beam from left to right (from 180 to 0 degrees at the beamformer). Therefore, the plot summarizes the whole sounding phase showing the sound power received in four specific DOAs (315º, 225º, 135º, 45º) plus the direct direction (0º) with the delay applied to the signal sent from the loudspeaker array.

The behavior of the sound in those DOAs is very important in order to know which delay should be used at the loudspeaker array during the sounding phase. All the angles of arrival except 0º, which is only shown to see the effect caused by the direct sound waves, correspond to the four channels explained in section 1.4: front right (315º), rear right (225º), rear left (135º) and front left (45º). Most of the times those angles are unreachable, but the selected DOAs are always the closest to them.

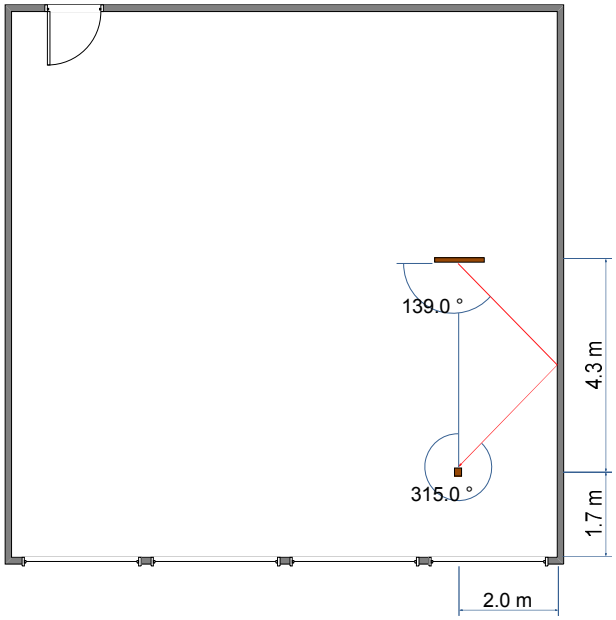



**Figure 6.3.** Imaginary trajectory of the sound on the laboratory when the sound projector transmits it with an angle of 139º.

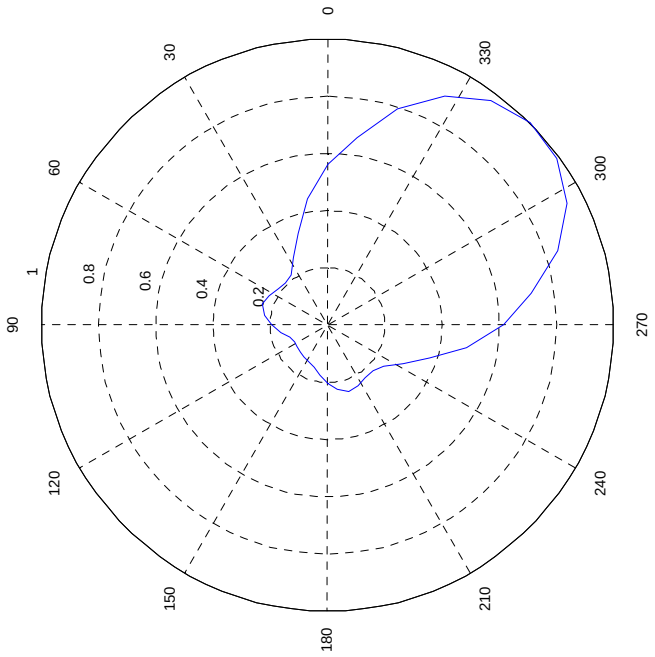



**Figure 6.4.** Radiation pattern received at the microphone constellation when the loudspeaker array is transmitting the signal with an angle of 139º (delay of -3.5 samples). It has been obtained from the CC technique.

A time-DOA plot is illustrated in Figure 6.2. This plot is obtained using the steps described in subsection 3.3.2.3. In this graphic, the received instantaneous sound power of all the soundings in a single long sweep is shown. Each direction of transmission of the sounding phase can be differentiated, being 19 in all. This plot reflects the behavior observed in Figure 6.1. The lowest received power is painted in dark blue and the colored areas which stand out of it show the

intensity and the angle from which they are obtained. Therefore, it can be noticed that when the sound projector directs the beam towards the left wall, which in this graphic corresponds to values of steering delay higher than 0.5 samples, no signal is received except a very low sound corresponding to the direct beam.

If a deeper analysis of the previous plots is done, it can be seen that for the direction of arrival of 315º (also the red line in Figure 6.1), the best power relation is obtained when a delay of -3.5 samples is applied at the loudspeaker array. This delay corresponds to a direction of transmission at the loudspeaker array of 139º.

Focusing on this direction of transmission, the radiation pattern received at the matrix of microphones is shown in Figure 6.4. This radiation pattern is obtained using the steps described in subsection 3.3.2.3. The main beam clearly arrives from an angle of 315º. Moreover, this confirms what can be observed in Figure 6.2. For the third direction of transmission (delay of -3.5 samples), one area is differentiated at 315º. This area contains the main part of its power between 270º and 360º. The value of the delay given to channel 1 by the SAM-30 is verified in MATLAB.

Now, for the DOA of 225º (the black line in Figure 6.1), the best delay is more difficult to recognize and has low power. However, in the sixth direction of transmission (delay of -2 samples) the black line stands out from the rest. This direction of transmission corresponds to an angle of 116º at the loudspeaker array. This is also the same delay given to channel 2 by the SAM-30.

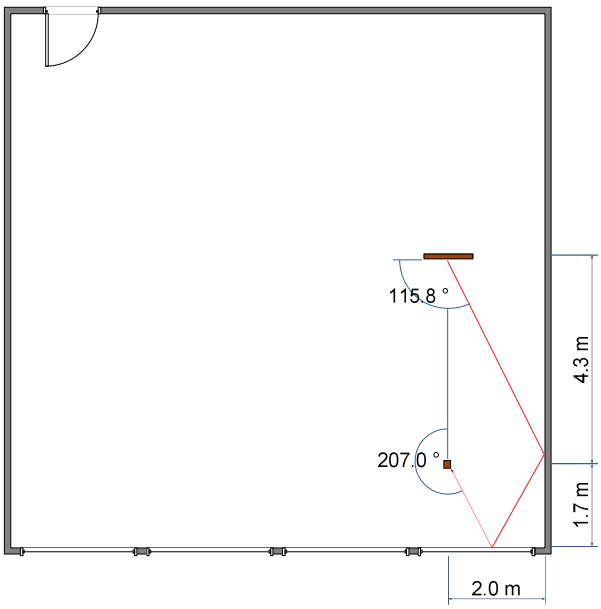



**Figure 6.5.** Imaginary trajectory of the sound on the laboratory when the sound projector transmits it with an angle of 116º.

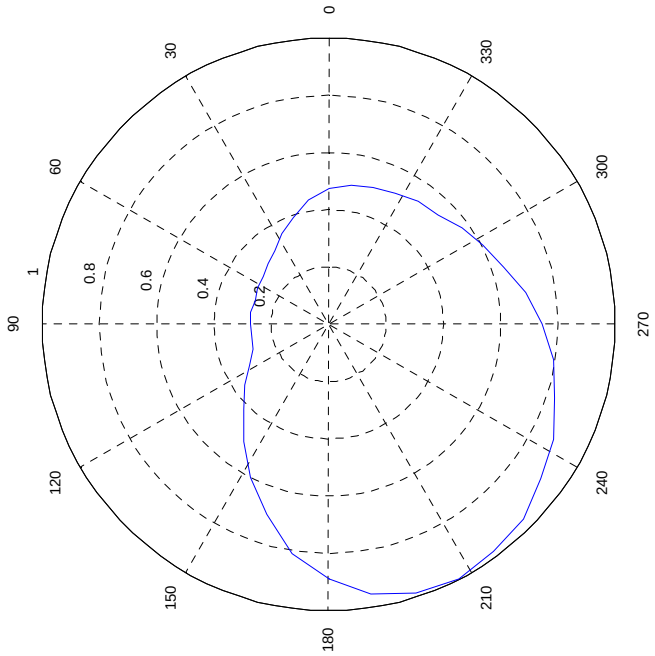



**Figure 6.6.** Radiation pattern received at the microphone constellation when the loudspeaker array is transmitting the signal with an angle of 116º (delay of -2 samples). It has been obtained from the CC technique.

The reason for the reception of this low sound power is the two reflections that the beam has to do. The first one happens on the right wall and the second on the window of the rear wall.

Although the sound power obtained at the sixth direction of transmission (delay of -2 samples) is low, the radiation pattern shows the behavior of the sound very well (Figure 6.6): it shows the sound waves arriving at the sensors from 207º. However, the radiation pattern also shows that its beamwidth is very wide which might affect to the DOA perception for human hearing. Looking at Figure 6.2, for a steering delay of -2 samples, the estimated instantaneous sound power is also low, but a sky-blue area can be noticed near to 210º.

The next point to analyze in Figure 6.1 and Figure 6.2 is the peak obtained when transmitting the signal with a delay of zero samples. However, this direction of transmission lacks interest because the signal is received at the position of the listener from the front (angle of 0º), and as it has been stated in section 1.4 the four directions to study are: front left/right and rear left/right.

Moreover, as it is explained at the beginning of this section, due to the position of the loudspeaker on the room, the radiation patterns and other graphics obtained from directions of transmission less than 90º are useless and are not analyzed. However, all the additional graphics can be found in Appendix D.

#### 6.1.2.2 Tests Performed in a Noisy Environment

In order to simulate a noisy environment, a new speaker was used. It was located at 80º of the microphones and it transmitted an additive Gaussian noise during the sounding phase. It was tested with different amplitude levels. The results shown in this subsection are given from five tests with a noise signal whose amplitude was 1.3 *dB* less than the true one.

The average values of the variables obtained from the five experiments are:

| Variable | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| `steering_delay` | 0 | 0 | 4.5 | 4.5 |
| `attenuation` | 0.7361 | 0.7361 | 0.9974 | 1.0000 |

**Table 6.2.** Average values of `steering_delay` and `attenuation` obtained from the CC technique for the four channels in a noisy environment.

Focusing on channel 1 and 2, a zero delay means that, during the projector phase, the loudspeaker array has to transmit the signal without doing any beamforming, so all the sound arrives at the position of the listener from the front. As it can be observed, the `steering_delay` and `attenuation` variables obtain completely wrong values due to the effect of the noisy source.

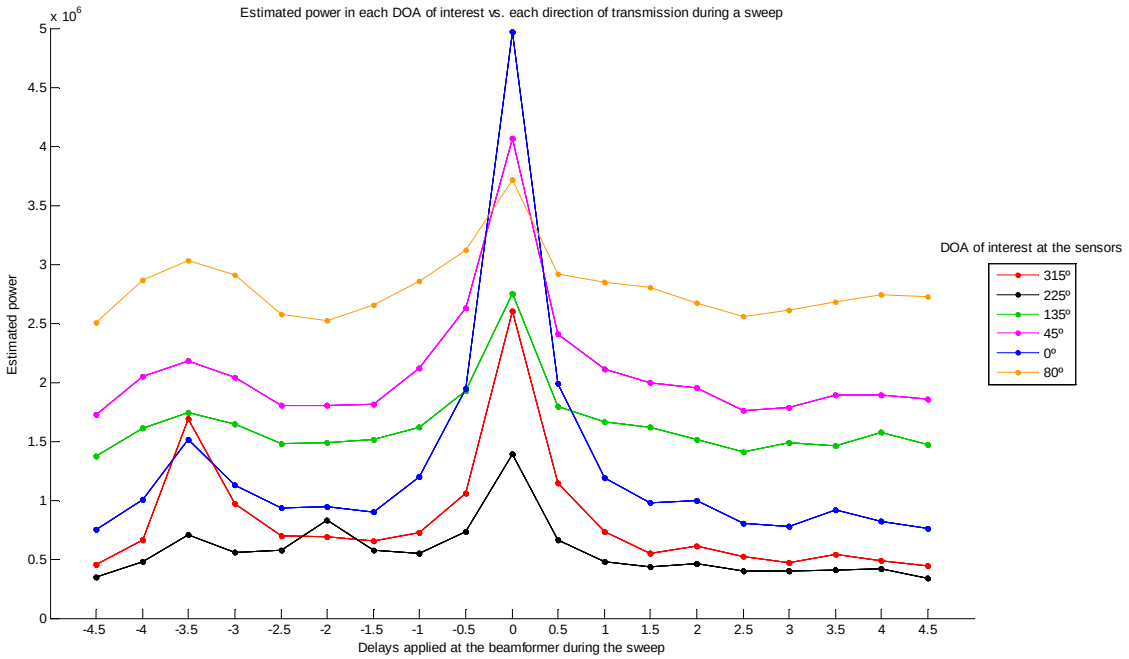



**Figure 6.7.** Estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) in a noisy environment. It has been obtained from the CC technique.

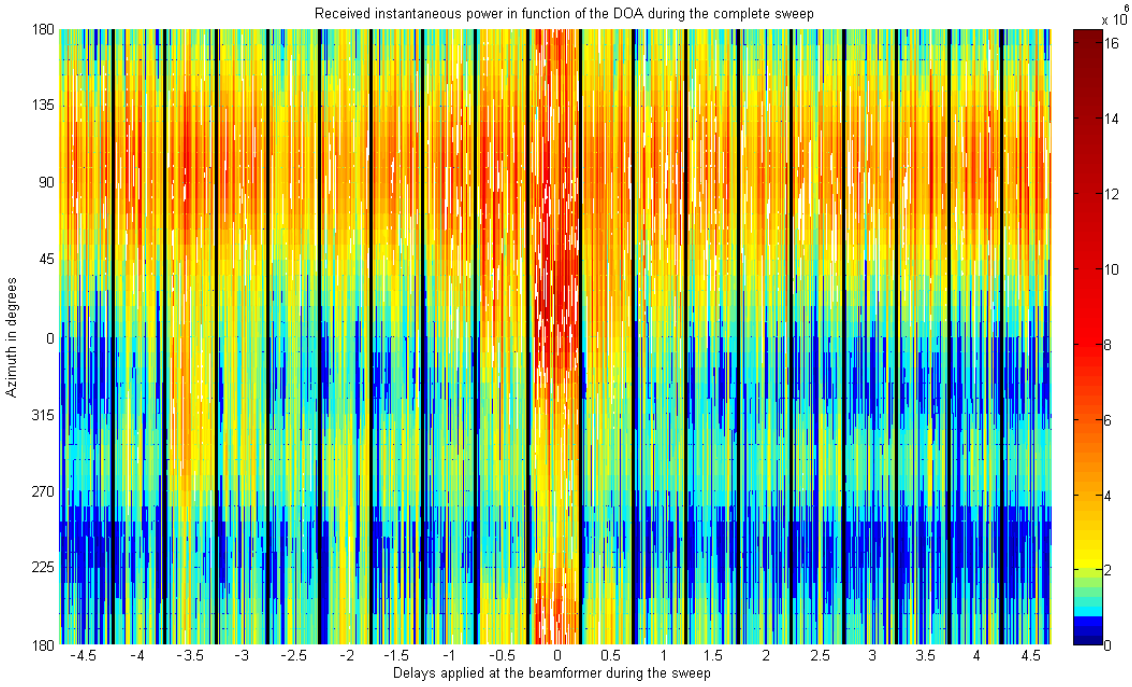



**Figure 6.8.** Time-DOA plot of the received instantaneous sound power during the sounding phase in a noisy environment. It has been obtained from the CC technique.

Comparing Figure 6.1 to Figure 6.7, an evident difference can be observed: the received sound power from some specific DOAs get on average higher values. This is caused by the interfering noise which modifies and distorts the analysis. Now, the peaks that stood out from the rest in the previous test are not so clear. The high level of this interfering signal hides them and makes it difficult to perform the analysis.

In Figure 6.7, an orange line has been added. It indicates the sound power received at the sensors from an angle of 80º (the position where the noisy source was located and from where the interfering signal is received) during the sounding phase. As it can be observed, the orange line has higher average values. Thus, it is closely related to the noisy source and signals its sound power. Further evidence of this is that the line almost maintains a constant level and has no strong fluctuations.

This hypothesis is also supported by Figure 6.8. There, a red-yellow area covers a range of DOAs from 60º to 140º during the 19 soundings. This area maintains almost the same intensity during the sounding phase which is interpreted as a result of the noisy source. As it can be observed, this pattern is totally disturbed by this signal.

Although the interfering noise is very loud (orange line in Figure 6.7 and red-yellow range in Figure 6.8), the first peak, corresponding to the third direction of transmission (delay of -3.5 samples), is still visible. Its radiation pattern (see Figure 6.4) suggests that the contribution of this interfering noise is important and it almost hides the signal totally.

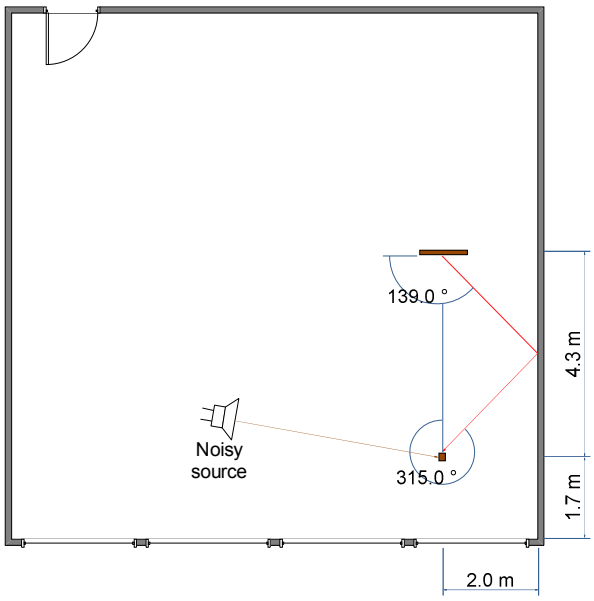



**Figure 6.9.** Imaginary trajectory of the sound on the laboratory when the sound projector transmits it with an angle of 139º. In this case, the noisy source, which is located at 80º from the sensors, is also transmitting.

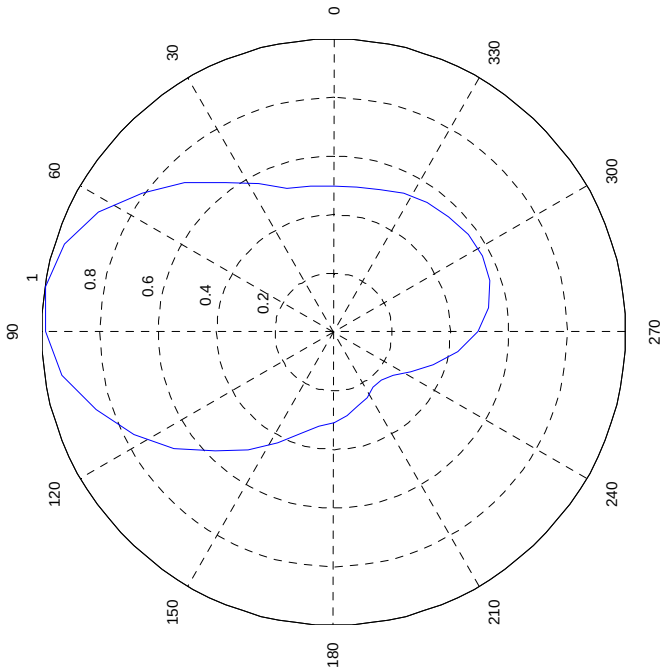



**Figure 6.10.** Radiation pattern received at the sensors with the interference of a noisy source when the loudspeaker array is transmitting the signal with an angle of 139º (delay of -3.5 samples). It has been obtained from the CC technique.

When comparing Figure 6.4 to Figure 6.10, there is a clear difference: another lobe appeared in the direction of the noisy source (80º) and the true one almost disappeared (300º). That fact makes the system believe that the signal sent by the loudspeaker array is arriving at the microphones from 80º, but this is the DOA of the interfering noise.

Having in mind Figure 6.6 and knowing that, in a quiet environment, the best delay for a DOA of 225º is -2 samples, it would be a good idea to look at the radiation pattern of that direction of transmission and compare it with the one on the previous chapter. Figure 6.12 is the equivalent in the current test. The first impression is that both figures are completely different. The fact is that the interfering signal totally masks the transmitted signal. The expected signal from an angle of 198º is too weak to appear in the radiation pattern and the only lobe belongs to the noisy source added in this test.

In those conditions, this method cannot determine the correct parameters. It will choose wrong DOAs and it will use wrong delays during the projector phase.

With these examples (Figure 6.10 and Figure 6.12), the lack of robustness of the CC technique is shown. The main problem of this method is that the transmitted signal is not used in the calculations at all. It only uses the received signals from the four microphones, which means (as theoretically shown in subsection 3.3.4) that if interfering and undesired signals appear during the sounding phase, the results and obtained parameters will be affected. Therefore, the CC technique must be run in a quiet environment in order to avoid being affected by an external and undesired noise.

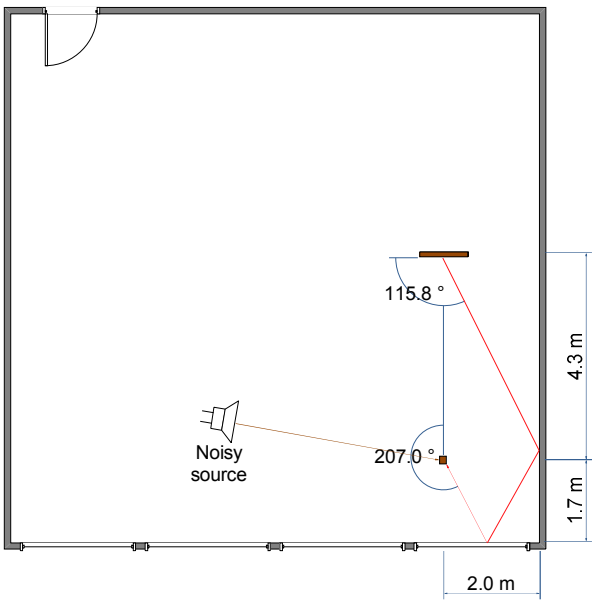



**Figure 6.11.** Imaginary trajectory of the sound on the laboratory when the sound projector transmits it with an angle of 116º. In this case, the noisy source, which is located at 80º from the sensors, is also transmitting.

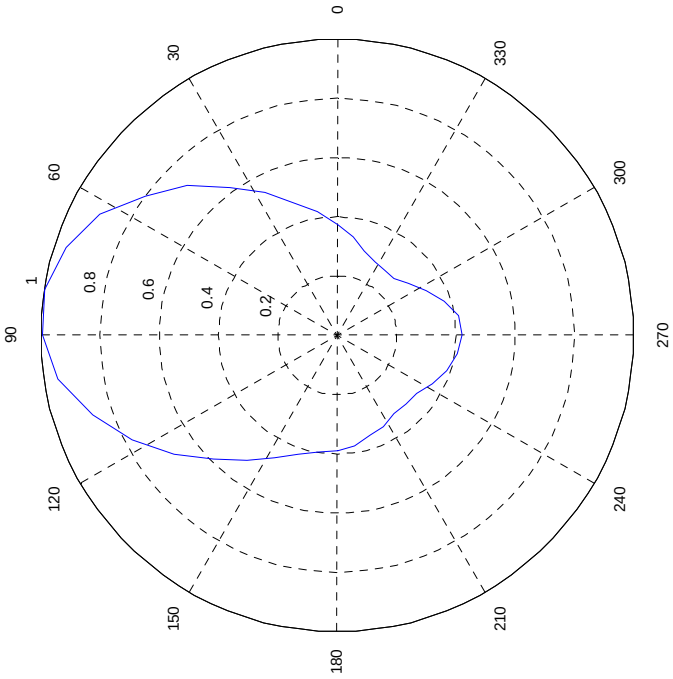



**Figure 6.12.** Radiation pattern received at the sensors with the interference of a noisy source when the loudspeaker array is transmitting the signal with an angle of 116º (delay of -2 samples). It has been obtained from the CC technique.

### 6.1.3 Results Obtained from the CSS Technique

The lack of robustness of the CC technique under noisy environments is solved by using the CSS technique, which is theoretically explained in subsection 3.3.3. It uses the transmitted and the four recorded signals in the calculations. This manner the external noises can be avoided and rejected, obtaining the correct parameters in order to run the projector phase. The CSS technique also allows the calculation of the propagation delay of the received signals.

In the following sections the results of the same tests, but now using the CSS technique, are shown and discussed.

#### 6.1.3.1 Tests Performed in a Quiet Environment

After running the ten experiments with the CSS technique, the average values obtained for channel 1 and 2 are shown in Table 6.3 (see Appendix C).

The new variable `transmit_delay` (see subsection 5.1) is obtained from the CSS method. It synchronizes the sound transmitted through the four channels at the position of the listener during the projector phase. It indicates the samples each channel has to be delayed in order to suppress the difference between propagation delays. In this case, sound from channel 2 must be transmitted first (without being delayed), then sound from channel 1 must be delayed 385 samples in order to arrive at the position of the listener simultaneously.

| Variable | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **`steering_delay`** | -3.5 | -2 |
| **`attenuation`** | 0.6492 | 1.0000 |
| **`transmit_delay`** | 391 | 0 |

**Table 6.3.** Average values of `steering_delay`, `attenuation` and `transmit_delay` obtained from the CSS technique for channels 1 and 2 in a quiet environment.

Comparing the values of `steering_delay` and `attenuation` obtained from the CSS technique (see Table 6.3) with the ones obtained from the CC method (see Table 6.1), it can be observed that they are very similar. This makes us think that, under a quiet environment, either method works properly.

Following the scheme developed in the previous chapters, Figure 6.13 shows a summary of the received sound power of some specific DOAs during the sounding phase. Looking at this figure, there are two points where the red and the black line (corresponding to an angle of arrival of 315º and 225º respectively) stand out from the rest.

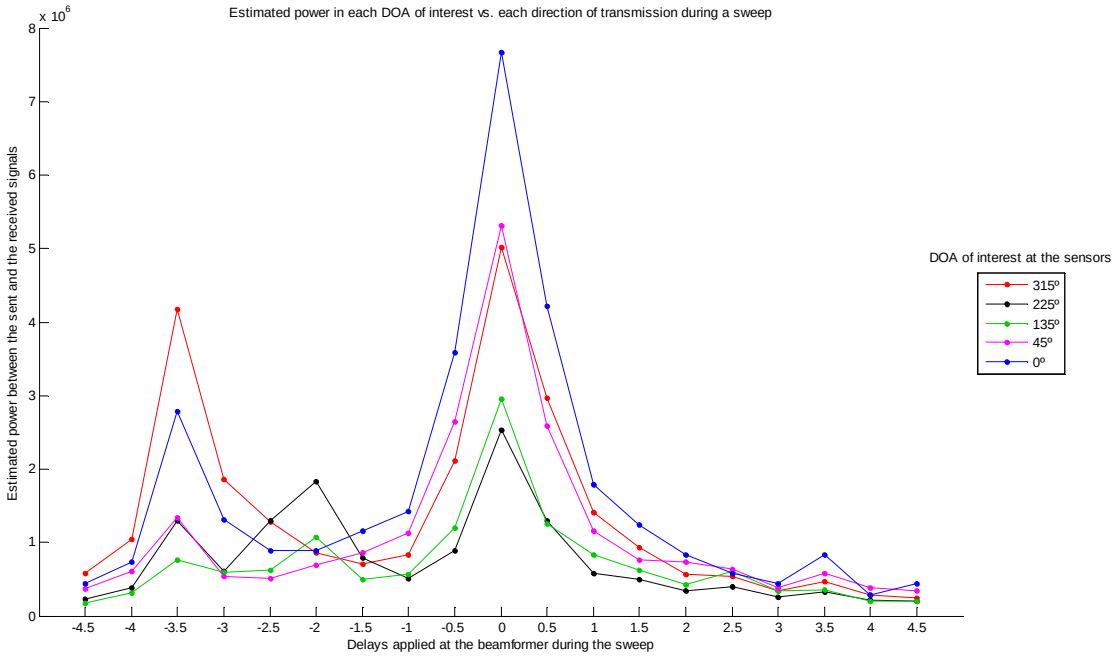



**Figure 6.13.** Estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) for a quiet environment. It has been obtained from the CSS technique.

For the red line, that high point is the peak at the third direction of transmission (delay of -3.5 samples). For the black line, that point is located at the sixth direction of transmission (delay of -2 samples). In those points, the angle of transmission at the loudspeaker array is 139º and 116º respectively. This agrees with the discussion from section 6.1.2.1.

The CSS technique (see subsection 3.3.3.1) allows a new plot to be added (Figure 6.14). This is a propagation delay-DOA plot obtained during the transmission of the signal in a specific direction. At the Y axis, the DOA is set; and at the X axis, the propagation delay is set. The red marks indicate when and from where the sound beams are received. The difference between this plot and the radiation pattern is that in this figure the moment when the sound waves from different reflections arrive at the microphones is shown. In the radiation pattern it cannot be determined which sound waves arrive first at the position of the sensors, nor the instant they arrive at.

The following propagation delay-DOA plot is obtained, from one of the five tests, when the loudspeaker array is transmitting the signal with a delay of -3.5 samples. As it can be seen, only one sound wave appears at 18.7 *ms*. The red area, which also indicates from where the signal is arriving, ranges from 270º to 15º, with the center located at 315º. Thus, it can be stated that the direct signal is rejected.

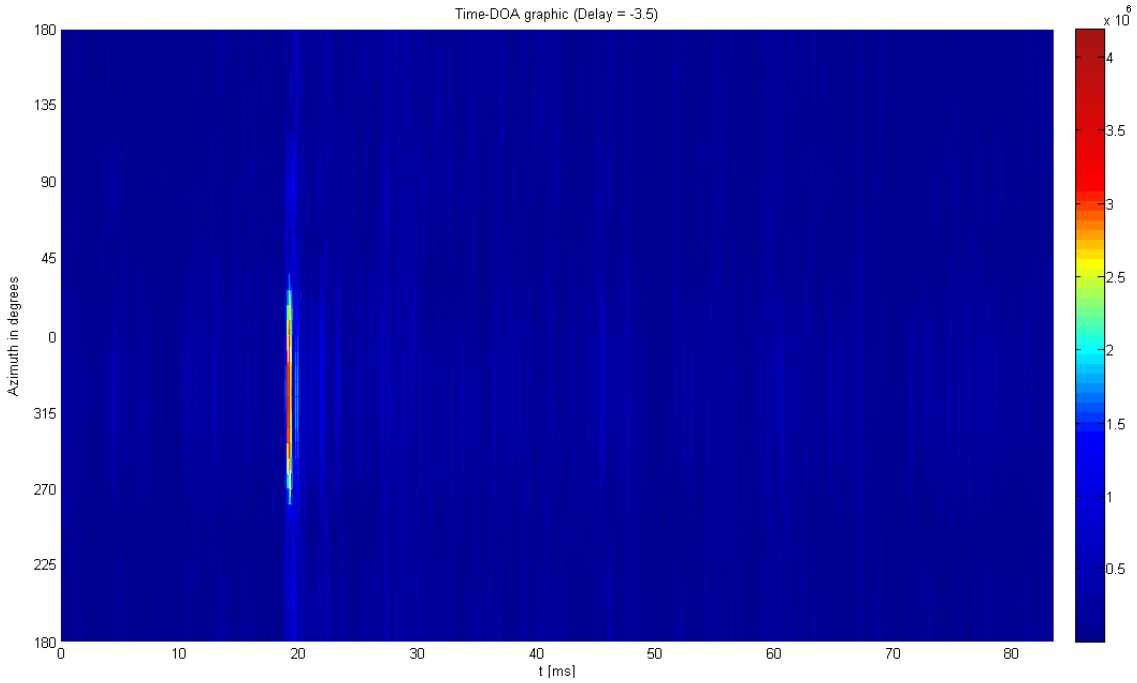



**Figure 6.14.** Propagation delay-DOA plot for a steering delay of -3.5 samples at the sound projector in a quiet environment. It has been obtained from the CSS technique.

Thus, knowing that the signal arrives at the position of the microphones 18.7 *ms* after starting the transmission, the covered distance can be calculated:

$$d_{315^o} = 18.7 \cdot 10^{-3} s \cdot 340 m/s = 6.4 m \tag{6.1}$$

In order to verify the good behavior of this method, the distance between the sound projector and the microphones should match up with the theoretical one (Figure 6.15). The results show that the signal arrives at the position of the listener from an angle of 315º, and it is also known that the loudspeaker array transmits the signal with an angle of 139º. The configuration is very close to an isosceles right-angled triangle and thus it can be approximated by the Pythagorean theorem:

$$d \simeq \sqrt{2 \cdot \left( \frac{d_{315^o}}{2} \right)^2} = 4.5 m \tag{6.2}$$

Due to the applied approximations, the value obtained in (6.2) differs slightly with respect to the theoretical one (section 2.1). However, as it was expected, the obtained values are similar enough to verify the correct performance of the equipment.

Also, the radiation pattern, which is shown in Figure 6.16, verifies what could be seen in the time-DOA plot above. This radiation pattern is obtained from the steps described in subsection 3.3.3.2. It shows a main lobe centered at 315º and no sound waves in any other direction.

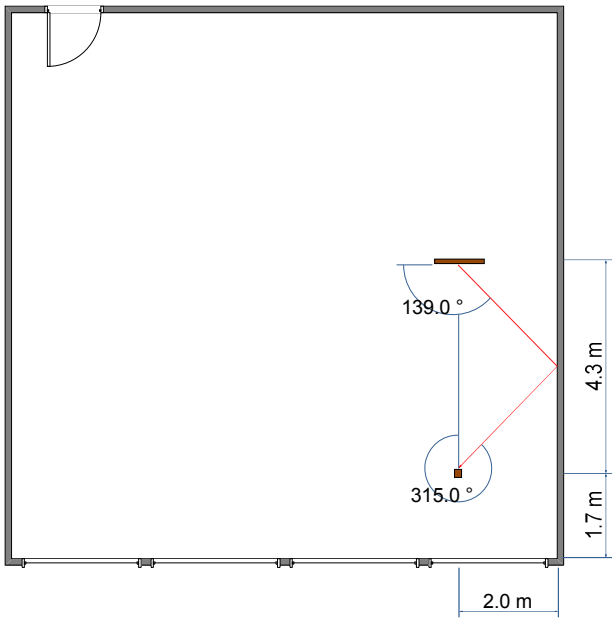



**Figure 6.15.** Imaginary trajectory of the sound on the laboratory when the sound projector transmits it with an angle of 139º.

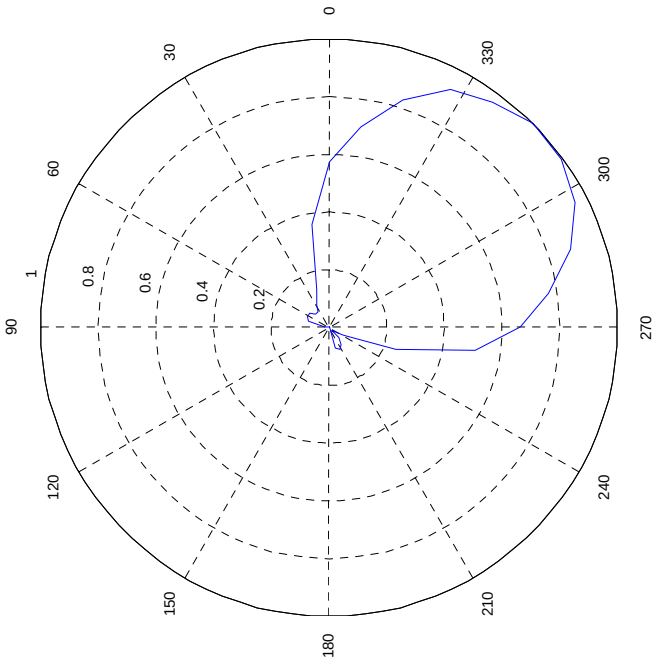



**Figure 6.16.** Radiation pattern received at the microphone constellation when the loudspeaker array is transmitting the signal with an angle of 139º (delay of -3.5 samples). It has been obtained from the CSS technique.

At this moment, it is time to analyze what happens when the signal is sent with a direction of transmission of 116º (delay of -2 samples). Figure 6.17 corresponds to its propagation delay-DOA plot. There are two areas to discuss: the first it is located at 13.1 *ms* and centered at 9º; the second is located at 27.1 *ms* and centered at 207º.

The first area is very low powered and corresponds to the direct beam. The distance covered by these first sound waves is calculated as follows:

$$d_{0^o} = 13.1 \cdot 10^{-3} s \cdot 340\, m/s = 4.5\, m \tag{6.3}$$

The second area corresponds to the beam which arrives from the right rear side of the microphones (207º). In this case, the distance covered by this beam is:

$$d_{207^o} = 27.4 \cdot 10^{-3} s \cdot 340\, m/s = 9.3\, m \tag{6.4}$$

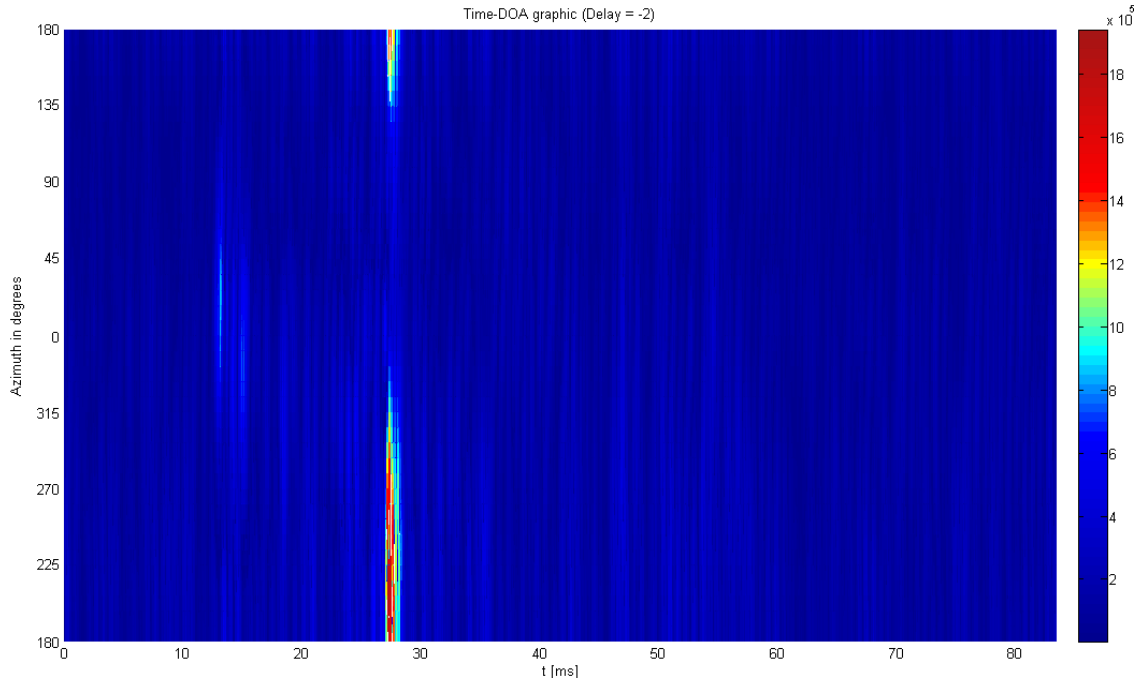



**Figure 6.17.** Propagation delay-DOA plot for a steering delay of -2 samples at the sound projector in a quiet environment. It has been obtained from the CSS technique.

The trajectory followed by these sound waves should be similar to the one shown in Figure 6.18. If this is true, $a_1$+$b_1$+$c_1$ should approximately score $d_{116^o}$.

$$a_1 = \frac{2\ m}{\sin\left(25.8^o\right)} = 4.6\ m; \qquad a_2 = \frac{2\ m}{\tan\left(25.8^o\right)} = 4.1\ m$$

$$b_2 = 6 - a_2; \qquad b_1 = \frac{b_2}{\cos\left(25.8^o\right)} = 2.1\ m; \qquad b_3 = b_2 \cdot \tan\left(25.8^o\right) = 0.9\ m$$

$$c_3 = 2 - b_3; \qquad c_1 = \frac{c_3}{\cos\left(25.8^o\right)} = 2.5\ m$$

So, the final value of the theoretical distance is:

$$d_{th,207^o} = a_1 + b_1 + c_1 = 9.2\ m \tag{6.5}$$

Then, the theoretical and the measured value are similar enough to determine that this trajectory is the one that the sound waves covered.

Figure 6.19 illustrates the radiation pattern of the current analysis (for a delay of -2 samples). It is very similar to Figure 6.12, which was obtained under the same conditions.

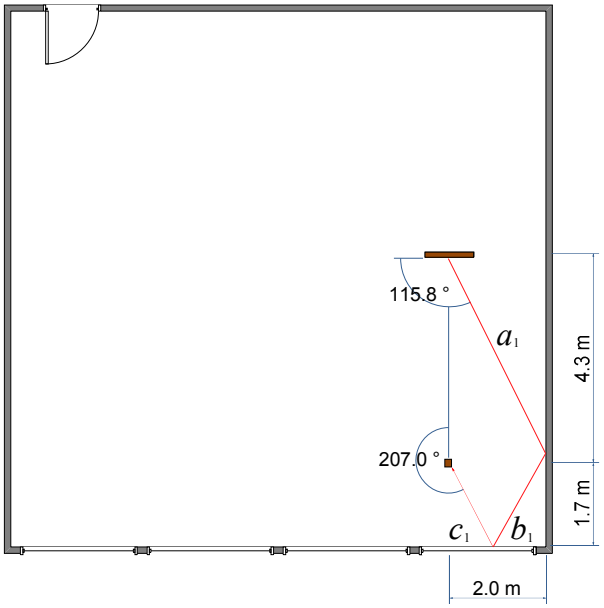



**Figure 6.18.** Imaginary trajectory of the sound on the laboratory when the sound projector transmits it with an angle of 116º.

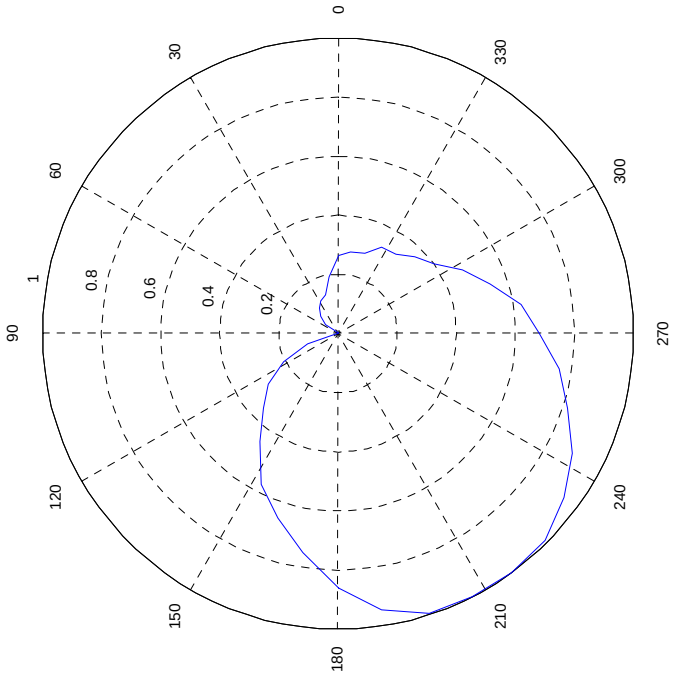



**Figure 6.19.** Radiation pattern received at the microphone constellation when the loudspeaker array is transmitting the signal with an angle of 116º (delay of -2 samples). It has been obtained from the CSS technique.

An interesting part of this analysis is to discover what happens with the transmitted signals steered towards the left wall. In fact, this method can contribute to understand it. Figure 6.20 is the propagation delay-DOA plot that belongs to the last direction of transmission (delay of 4.5 samples) and it shows its behavior very well.

First, the signal is transmitted from the sound projector with an angle of 11º, it reflects on the left wall and reaches the sensors. Then, it follows to the right wall, reflects on it and reaches the microphones again. Meanwhile, a beam goes directly to them from the loudspeaker array.

In the figure, the red area with the center at 9º and located at 14.5 *ms* corresponds to the direct beam; the area with the center at 72º and located at 56.3 *ms* corresponds to the main beam after the first reflection on the left wall; and, finally, the area centered at 279º and located at 68.0 *ms* corresponds to the main beam after the second reflection on the right wall.

As it can be observed, the main and the direct beam have almost the same power. This is because of the reflections and the length between the loudspeaker array and the left wall of the room. Doing some calculations, the covered distance can be obtained:

- Distance covered by the main beam until the first time it reaches the microphones.

$$d_{72^o} = 56.3 \cdot 10^{-3} s \cdot 340\, m/s = 19.1\, m \tag{6.6}$$

- Distance covered by the main beam between the first and the second time it reaches the sensors.

$$d_{279^o} = \left(68.0 \cdot 10^{-3} - 56.3 \cdot 10^{-3}\right) s \cdot 340\, m/s = 4.0\, m \tag{6.7}$$

- Distance covered by the direct beam.

$$d_{9^o} = 14.5 \cdot 10^{-3} s \cdot 340\, m/s = 4.9\, m \tag{6.8}$$

All these distances meet the dimensions of the room: $d_{72^o}$ and $d_{279^o}$ are approximately twice the distance from the loudspeaker array to the left and to the right walls respectively, and $d_{0^o}$ is a little bit more than the space between the sound projector and the microphones.

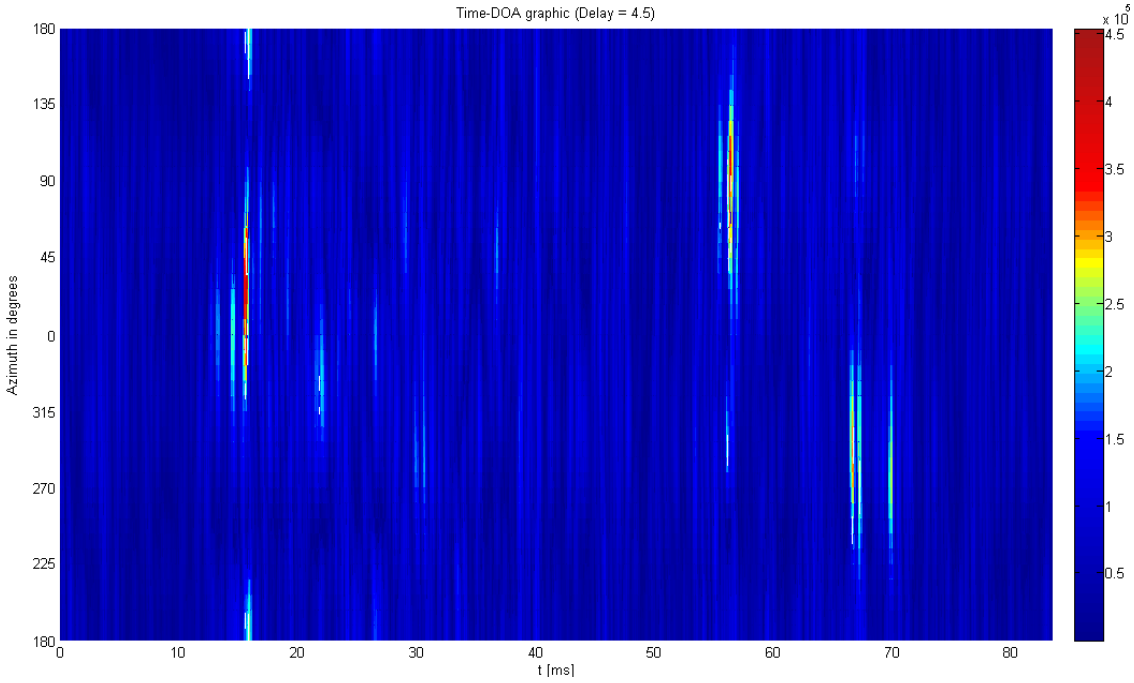



**Figure 6.20.** Propagation delay-DOA plot for a steering delay of 4.5 samples at the sound projector in a quiet environment. It has been obtained from the CSS technique.

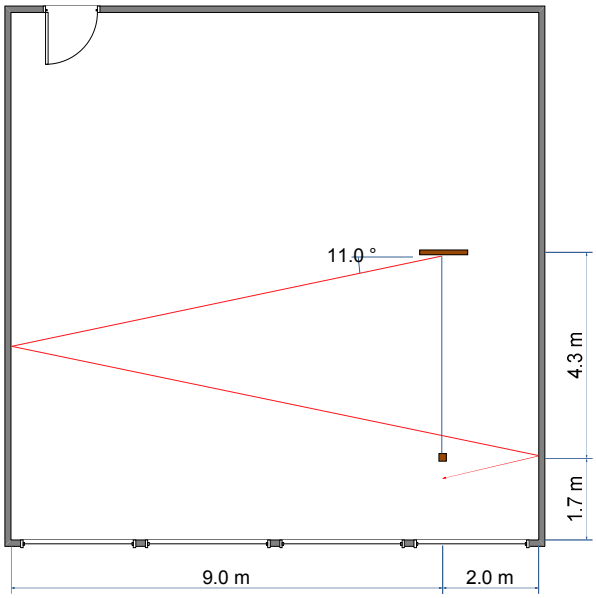



**Figure 6.21.** Imaginary trajectory of the sound on the laboratory when the sound projector transmits it with an angle of 11º.

### 6.1.3.2 Tests Performed in a Noisy Environment

The last situation to analyze is CSS technique working in a noisy environment. In this case, the measurements are not distorted by the noisy source located at 80º from the position of the listener. It only affects the results in directions of transmission in which the received sound power is very low, as it can be seen in Appendix D.

The following graphs show that the contribution of the noisy source is rejected at the directions of interest, and the obtained results are very similar to the ones shown in the previous chapter. Due to the minimum differences between them, this case will not be analyzed specifically because the conclusions are the same.

The values of the variables given by MATLAB after running the analysis of these recordings are the following. Channels 3 and 4 are not shown.

| Variable | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| `steering_delay` | -3.5 | -2 |
| `attenuation` | 0.6568 | 1.0000 |
| `delay_samples` | 389 | 0 |

**Table 6.4.** Average values of `steering_delay`, `attenuation` and `transmit_delay` obtained from the CSS technique for channels 1 and 2 in a noisy environment.

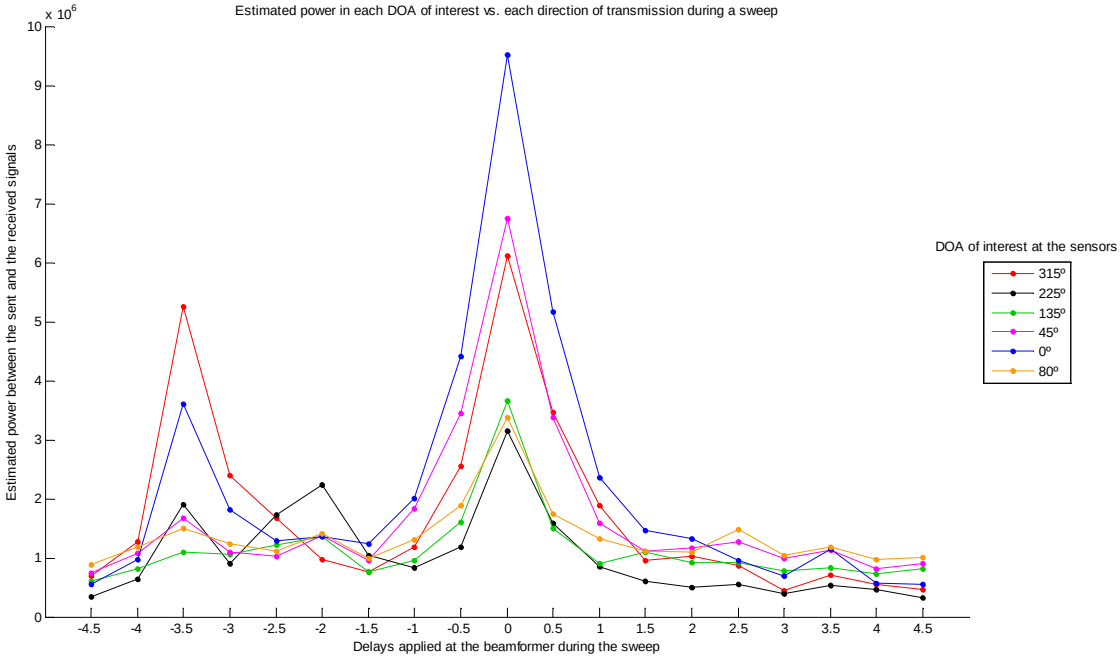



**Figure 6.22.** Estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) for a noisy environment. It has been obtained from the CSS technique.

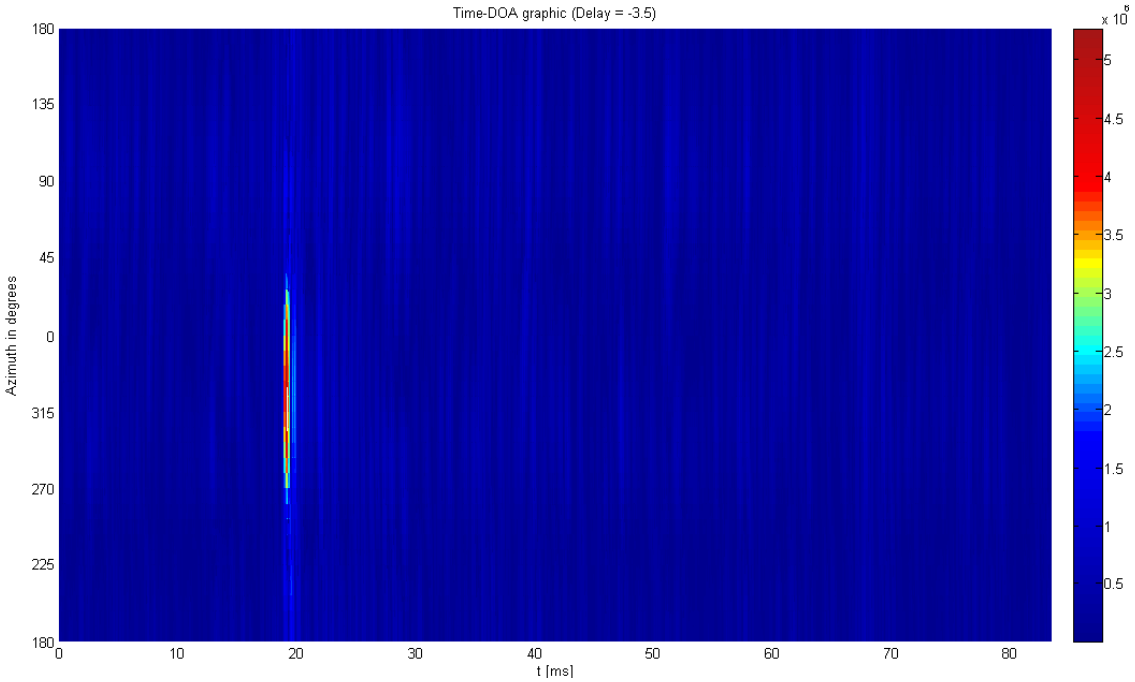



**Figure 6.23.** Propagation delay-DOA plot for a steering delay of -3.5 samples at the sound projector in a noisy environment. It has been obtained from the CSS technique.

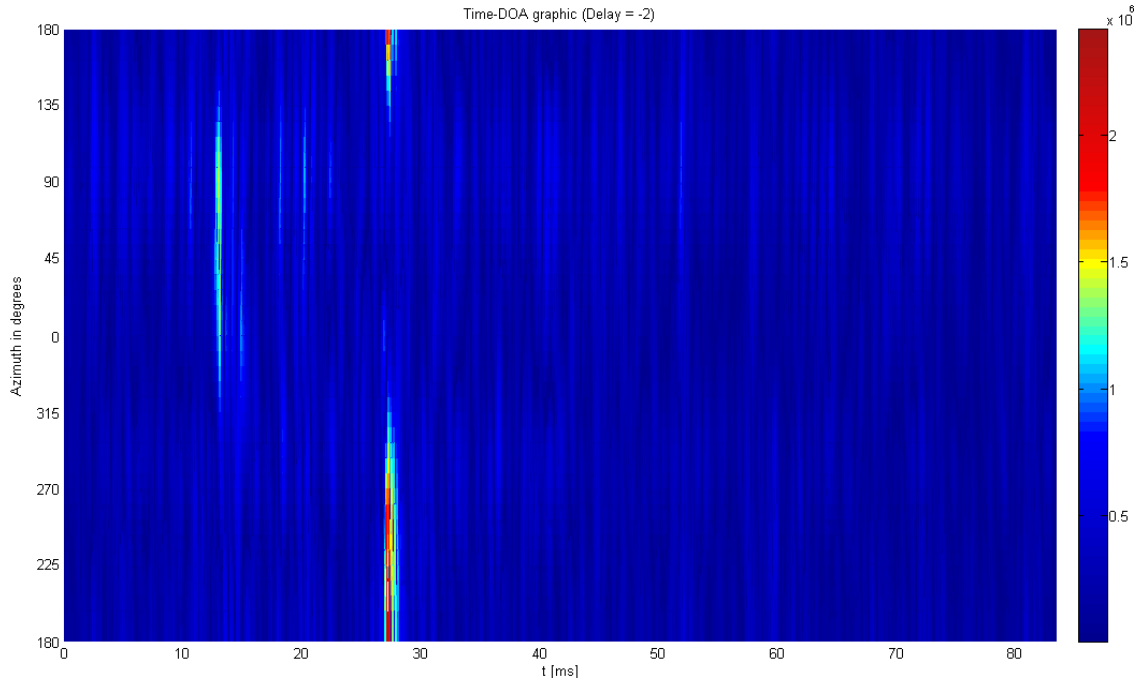



**Figure 6.24.** Propagation delay-DOA plot for a steering delay of -2 samples at the sound projector in a noisy environment. It has been obtained from the CSS technique.

### 6.1.4 Comparison of Both Techniques

In the previous subsections, the results obtained from two environmental situations have been shown:

- Results of ten tests performed in a quiet environment using the CC (subsection 6.1.2.1) and the CSS (subsection 6.1.3.1) techniques.
- Results of five tests performed in a noisy environment using the CC (subsection 6.1.2.2) and the CSS (subsection 6.1.3.2) techniques.

For the described laboratory room and position of the loudspeaker array and microphones, the analyses made in sections 6.1.2.1, 6.1.3.1 and 6.1.3.2 showed the same values of the `steering_delay` and `attenuation` variables (only front/rear right channels are taken into account; it has been already explained the problem of deciding the front/rear left channels).

The results obtained from the analyses of the CC and CSS techniques in a quiet environment are very similar. Under these external conditions, the CC and CSS techniques seem to work properly. The problems appear when another sound source is introduced in the room and it transmits an interfering signal which is also recorded on the sensors.

The CC method cannot remove this interfering signal and if it is powerful enough this technique begins to obtain erroneous parameters. This behavior has analyzed theoretically in section 3.3.4.1 and in practice in 6.1.2.2: the contribution of the interfering signal cannot be removed and depending on its sound power it could hide the signal transmitted by the loudspeaker array. Therefore, the system is not able to avoid external sounds by using the CC method.

This contrasts with the results shown in section 6.1.3.2. There, the interfering signal is rejected and the correct parameters are obtained. In this case, the CSS technique utilizes the sent signal. Furthermore, this method is able to determine the propagation delay of the each channel, so the system can synchronize the signals transmitted through each of them in order to make them arrive at the position of the listener at the same time.

The program implemented in the SAM-30 only performs the CC method. Therefore it may deliver wrong performance in a noisy environment. In order to avoid it, the signal played by the sound projector was set to be loud enough to reject speech and other noises or whistles.

## 6.2 Subjective Impressions

The results obtained from the SAM-30 and MATLAB, were also verified by a real human listener.

The session was divided in two parts. The first part was devoted to listening to the sound transmitted in each direction during the sounding phase. The second part was focused on running the projector phase and observing the effect produced on the listener. These parts were repeated several times and with different listeners.

In the first part, the sounding phase was run but in this case a listener replaced the microphones. As it happened in the objective results, the listeners could not recognize any direction of arrival when the loudspeaker array steered the sound waves towards the left wall. On the other hand, the sound waves that were steered towards the right wall produced good perception of the DOA. The direction of transmission of 139º (delay of -3.5) delivered the best perception of sound coming from the front right side (channel 1). However, the choice of the rear right side (channel 2) was not easy. Although the objective results showed that the direction of transmission of 116º (delay of -2) was the best for this channel, the listeners could recognize, albeit with difficulty, the origin of the sound.

In the second part, the projector phase was run with the parameters obtained from SAM-30. Despite the loss of sound perception from the back of the listener and the knowledge that only channels 1 and 2 had correct parameters, when the right and rear channels were played at the same time they produced a slight surround effect around the listener.

Another point to discuss is the effect produced by the low-pass filtering. Once it was implemented, it was observed that frequencies higher than 6 *kHz* were rejected. However, the use of the filtering caused an almost imperceptible distortion. On the other hand, without the use of the filtering, the alias sound beams (see subsection 3.5) were perceived for frequencies higher than 6 *kHz*. After comparing the advantages and drawbacks of implementing the low-pass filtering, it was decided to maintain it.

# Chapter 7

# Conclusions and Future Work

This chapter is divided in two sections: the first, Conclusions, contains the main findings of this work; the second, Future Work, suggests possible system improvements.

## 7.1 Conclusions

The loudspeaker array can only have good control of the sound beam in the frequency range from 1.7 *kHz* to 5 *kHz*. This limitation is caused by the dimensions of the sound projector and the distance between speakers. However, it was found experimentally that for this specific sound projector, frequencies higher than 5 *kHz* affect the perception of the direction of arrival in a larger magnitude than frequencies lower than 1.7 *kHz*. Thus, a low-pass filtering was needed.

During the sounding phase, only the frequency range between 1.7 *kHz* and 3.5 *kHz* can be used because the microphone constellation restricts the maximum frequency in order to avoid aliasing on the received signals.

Although it was observed that nineteen different directions of transmission are enough to find a direction of transmission for each channel, it would be better to upgrade the program and introduce more directions of transmission so as to have more options and intermediate values.

Objective results showed that the CC technique works quite well in a quiet environment; however, it loses all its effectiveness in a noisy environment. On the other hand, the CSS technique showed to be much more robust under these conditions, since the contribution of the external noise was removed. Moreover, the CSS technique allows the DSP to synchronize the sound transmitted by the four channels at the position of the listener.

Subjective experiments showed that the right front and rear channels gave a very good insight into the DOA of sound, while the left front and rear channels did not allow this to be

determined. Nevertheless, when all of the channels were played simultaneously, the listener experienced a surround sound effect. The subjective results reflect only the particular case of the location of the system in this room. This can be accounted to the fact that the left wall is more distant as illustrated in Figure 2.1.

## 7.2 Future Work

Two areas for future work emerge: possible changes on the hardware specification and improvements related to the software implementation.

### 7.2.1 Possible Changes on the Hardware Specification

As explained in section 3.5.1, the physical configuration of the loudspeaker array restricts the available sound frequency range. In order to have good control of the beam and reject the possible appearance of a grating lobe this range is limited to 1.7 - 5 *kHz*.

Thus, if more loudspeakers were included at the sound projector and the distance between them was reduced, no grating lobes would appear and the whole range of frequencies would be available.

Also, the use of a more complex configuration for the sensors with the addition of more microphones would increase the precision of the results and the range of frequencies available for the tests.

### 7.2.2 Improvements Related to the Software Implementation

One of the main restrictions of the program is the limited number of available angles to direct the sound waves. In fact, there are 19 possible directions to steer the beam. This number is limited by the interpolation factor of the signal. If the interpolation factor was increased, the possible directions would increase as well.

On the other hand, the order $19^{th}$-order low-pass FIR filter implemented on the Projector_phase_48kHz.pjt spends a lot of DSP cycles and does not allow the use of a higher interpolation factor. The filtering was carried out in order to reject the transmission of frequencies higher than 5 *kHz*. So, if the hardware specification was changed and the loudspeaker array permitted to transmit on the whole range of frequencies available for the speakers, the low-pass filter would not be needed.

Finally, another improvement would be to implement the CSS technique in the SAM-30 application. It was performed in MATLAB obtaining better and more precise results than the CC technique, especially under noisy environments.

# Bibliography


[1] Home Theater: Home Theater in a Box Systems. [Online]. http://www.hometheatermag.com/hometheaterinabox/

[2] A. Cardama Aznar, et al., *Antenas*, Second edition ed. Barcelona: Edicions UPC, 2002.

[3] H. L. Van Trees, *Optimum array processing*. John Wiley & Sons, Inc., 2002.

[4] M. Á. Lagunas, "Array Proceesing I: Conformación y estimación de dirección de llegada," UPC Notebook, 2007.

[5] L. J. Griffiths and C. W. Jim, "An Alternative Approach to Linearly Constrained Adaptive Beamforming," *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 1, pp. 27-34, Jan. 1982.

[6] H. Krim and M. Viberg, "Sensor Array Signal Processing: Two Decades Later," 1995.

[7] U.-H. Kim, J. Kim, D. Kim, H. Kim, and B.-J. You, "Speaker Localization Using the TDOA-based Feature Matrix for a Humanoid Robot," *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication*, Aug. 2008.

[8] K. Varma, "Time-Delay-Estimate Based Direction-of-Arrival Estimation for Speech in Reverberant Environments," 2002.

[9] M. Brandstein and W. Darren, *Microphone Arrays*. Berlin: Springer, 2001.

[10] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin: Springer, 2008.

[11] C. H. Knapp and G. C. Carter, "The generalized Correlation Method for Estimation of time Delay," *IEEE Transactions on Acoustic, Speech and Signal Processing*, vol. 4, pp. 320-

327, Aug. 1976.

[12] G. C. Carter, A. H. Nuttall, and P. G. Cable, "The smoothed coherence transform," *IEEE Proceding letters*, vol. 61, pp. 1497-1498, Oct. 1973.

[13] A. Brutti, M. Omologo, and P. Svaizer, "Comparison between different sound source localization techniques based on real data collection.," *IEEE Hands-Free Speech Communication and Microphone Arrays*, pp. 69-72, May 2008.

[14] M. Homologo and P. Svaizer, "Use of the crosspower-spectrum phase in acoustic event location," *IEEE Transactions on speech and audio processing*, vol. 5, no. 3, May 1997.

[15] K. Varma, T. Ikuma, and A. A. Beex, "Robust TDE-based DOA estimation for compact audio arrays," *IEEE*, 2002.

[16] M. Rafiqul Islam and I. A. H. Adam, "Performance study of direction of arrival (DOA) estimation algorithms for linear array antenna," *IEEE International Conference on Signal Processing Systems*, pp. 268-271, 2009.

[17] (2008) Educational DSP, LLC. [Online]. http://www.educationaldsp.com/

[18] N. Otsu, "A Threshold Selection Method for Gray Level Histograms," *IEEE Transactions on System, Man and Cybernetics*, Jan. 1979.

[19] M. Sezgin and B. Sankur, "Survey over image thresholding techniques," *Journal of Electronic Imaging*, vol. 13, no. 1, p. 146–165, Jan. 2004.

[20] J. Dmochowski, J. Benesty, and S. Affes, "Direction of Arrival Estimation Using the Parameterized Spatial Correlation Matrix," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 4, May 2007.

[21] J. B. Mariño, F. Vallverdú, J. A. Rodríguez, and A. Moreno, *Tratamiento digital de la señal*, Third edition ed. Barcelona: Edicions UPC, 1999.

[22] J. G. Proakis and D. G. Manolakis, *Digital signal processing*, Fourth edition ed. Pearson Prentice Hall, 2007.

[23] A. K. Bhattacharyya, *Phased Array Antennas*. John Wiley & Sons, 2006.

[24] J. Gonzalvo, J. A. Morán, and J. Melenchón, "Detección del ángulo de llegada con un array microfónico," URL.

[25] S. Ries, "Digital time-delay beamforming with interpolated signals," 2004.

[26] R. Pico, J. Ramis, J. Redondo, and J. Alba, "Medida de la función de transferencia de un maniquí acústico," EPSG.

[27] A. Mínguez and M. Recuero, "Moldeado espectral de ruido acústico," TecniAcústica, 2000.

[28] J. L. Mather, A. C. Fairhead, and M. P. Warden, "Implications of adaptive cancellation array processing for design of a space-based surveillance radar," Ministry of Defence, 1991.

[29] H. Mizoguchi, Y. Tamai, K. Shinoda, S. Kagami, and K. Nagashima, "Visually Steerable Sound Beam Forming System based on Face Tracking and Speaker Array," *IEEE Proceedings of the 17th International Conference on Pattern Recognition*, 2004.

[30] D. Bonilla-Hernández, D. H. Covarrubias-Rosales, and J. G. Arceo-Olague, "Mejora de la Resolución de Fuentes en Campo Cercano por Medio del Estimador UML," *IEEE LATIN AMERICA TRANSACTIONS*, vol. 4, no. 6, Dec. 2006.

[31] C. Peng, H. Chao-huan, M. Xiao-Chuan, L. Yi-Hui, and Y. Jun, "New wideband DOA estimation algorithm for planar arrays," *Proceedings of 2007 International Symposium on Intelligent Signal Processing and Communication Systems*, Nov. 2007.

[32] A. Campos, et al., "The sound projector," KTH, 2007.

[33] B. Pueo, J. Escolano, and M. Cobos, "Control de la directividad en arrays lineales de altavoces," Ministerio de Ciencia y Tecnología.

[34] K. S. Tan, W. S. Gan, J. Yang, and M. H. Er, "An efficient digital beamsteering system for difference frequency in parametric array," *IEEE*, 2004.

[35] Texas Instruments, "TMS320C67x DSP Library Programmer's Reference Guide," 2006.

[36] A. Abad, "A multi-microphone approach to speech processing in a smart-room environment," PhD Thesis, UPC, 2007.

[37] VECO Vansonic Enterprise, "Loudspeaker model 28KC08-1-A datasheet".

[38] J. Benesty, J. Chen, and Y. Huang, "Time-delay estimation via linear interpolation and cross-correlation," *IEEE Transactions on speech and audio processing*, vol. 12, no. 5, Sep. 2004.

[39] M. Horvat, H. Domitrovic, and A. Petosic, "Linear arrays with polynomial phase distribution," *Proceedings Electronics in Marine. 46th International Symposium*, Jun. 2004.

[40] V. N. Christofilakisa, P. Kostarakis, A. A. Alexandridis, F. Lazarakis, and K. Dangakis, "Increasing direction-of-transmission resolution in digital time-delay beamformers," *AEU - International Journal of Electronics and Communications*, vol. 62, no. 1, pp. 59-56, Jan. 2008.

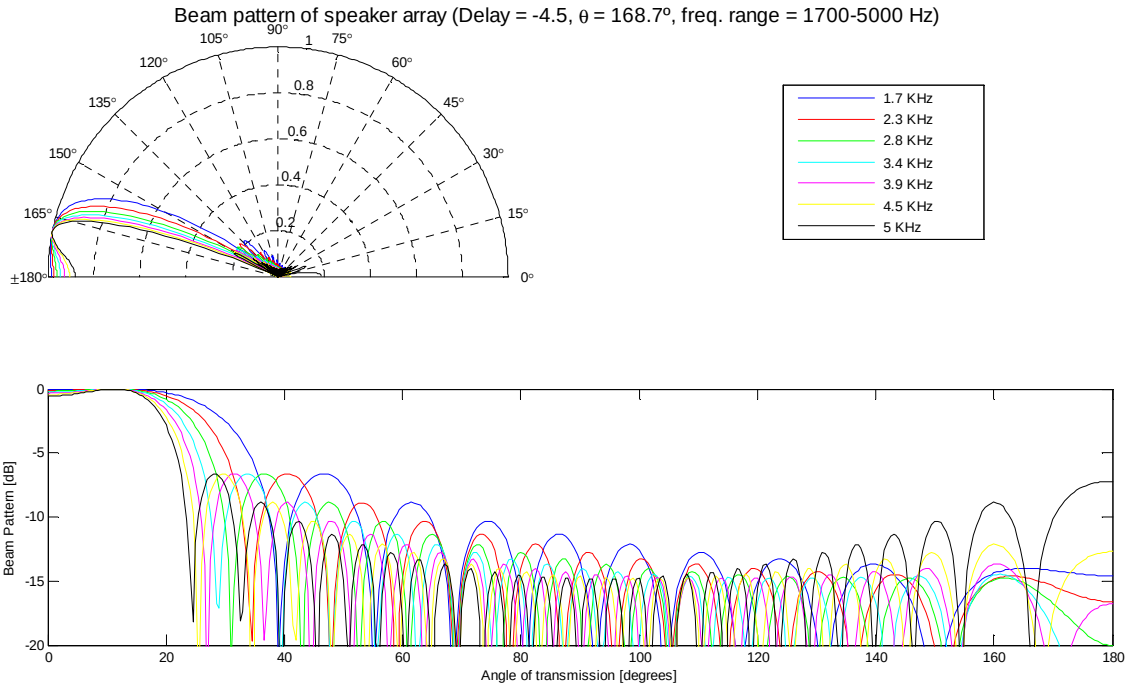
[41] S. Chawla, "Yamaha YSP-1 Sound Projector Produces Virtual Surround Sound with One Speaker Enclosure," *Secrets of Home Theater & High Fidelity*, vol. XII, no. 4, Oct. 2005.

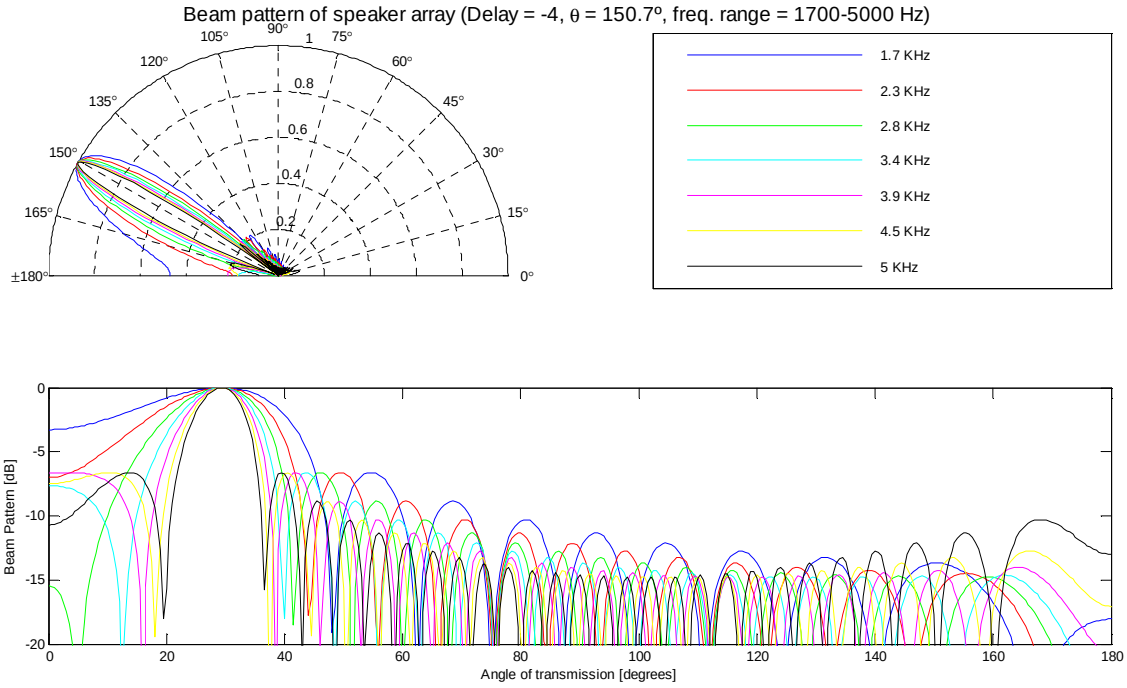
[42] Yamaha Electronics. [Online]. http://www.yamaha.com/yec/soundprojectors/

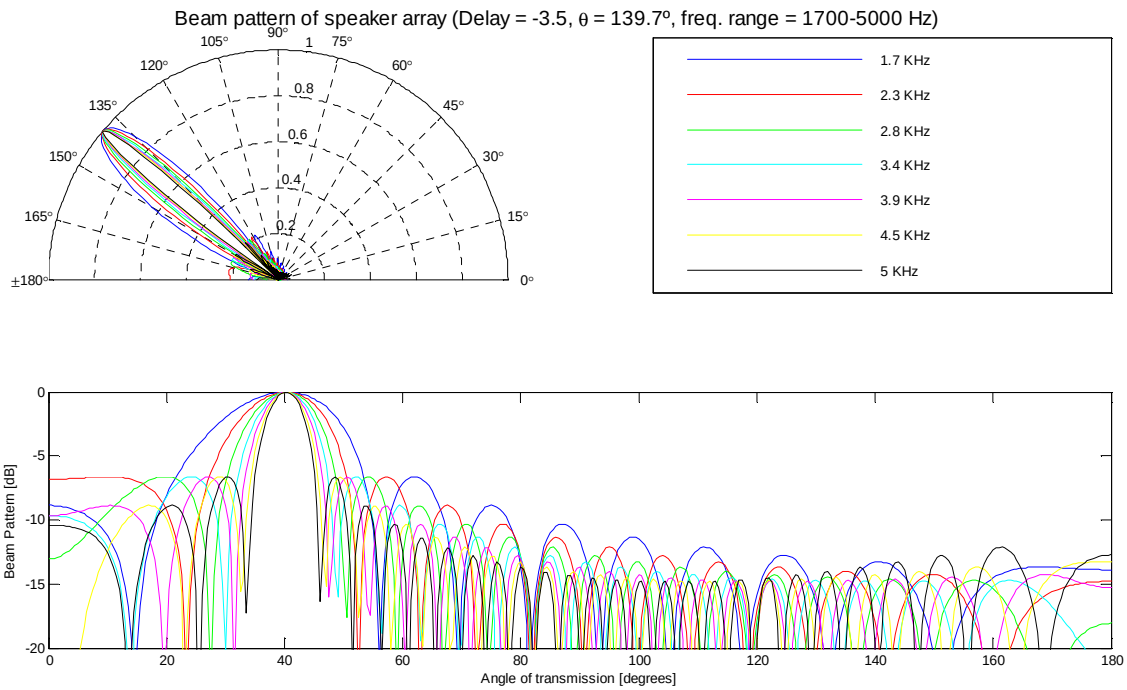
# Appendix A

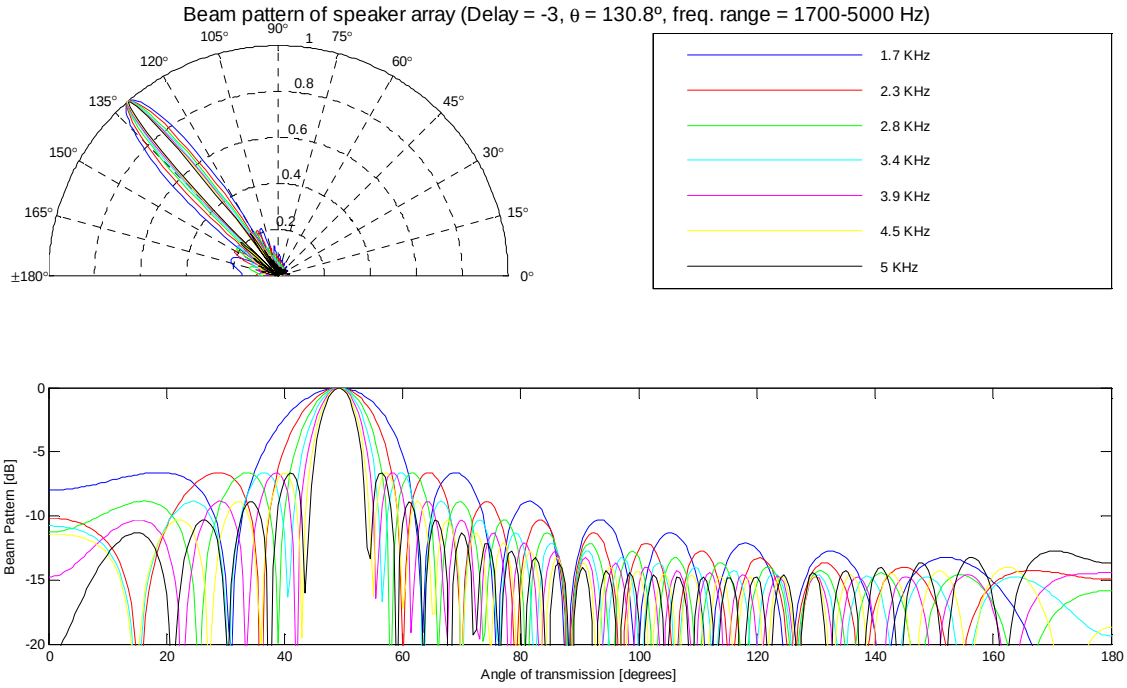
# Possible directions of transmission

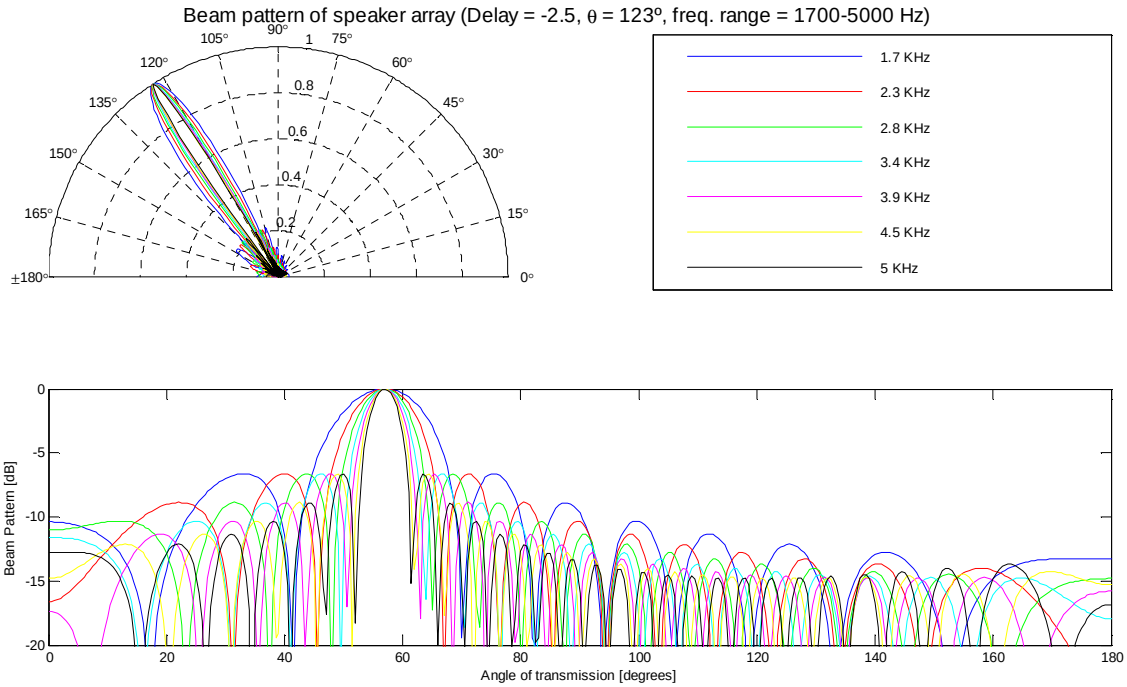
This appendix shows the 19 directions of transmission in which the loudspeaker array steers the beam. Due to their symmetry only the first ten (the ones corresponding to the left part of the radiation pattern) are shown. The figures show the Array Factor in linear (semicircle) and logarithmic (rectangle) scale for certain frequencies.

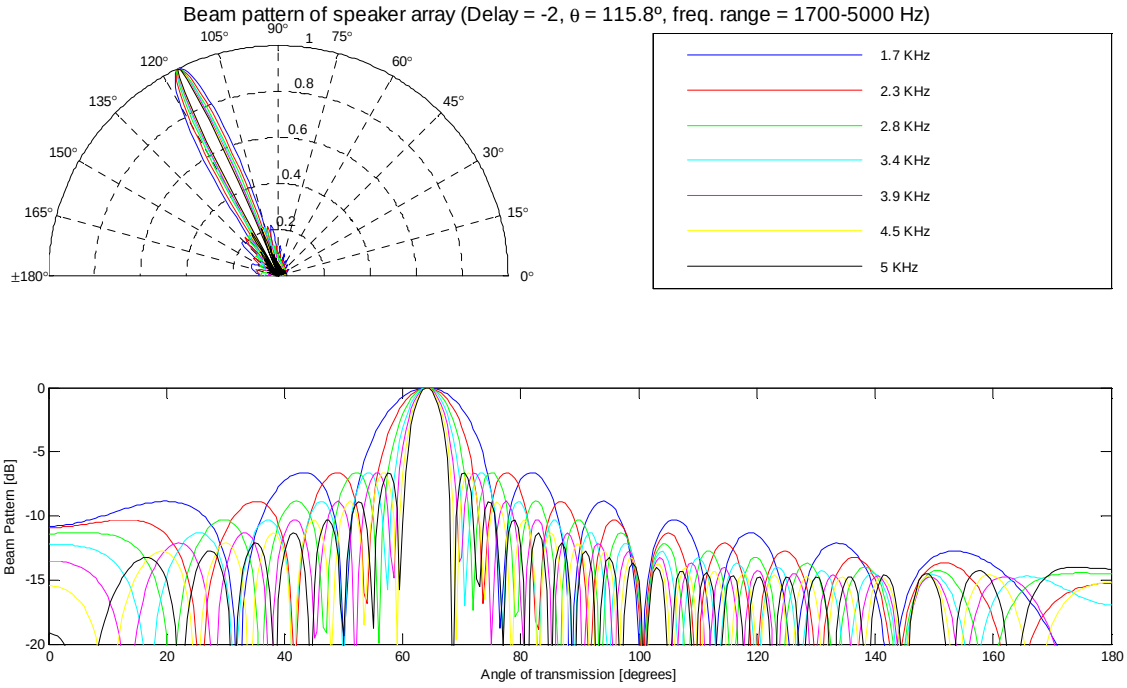


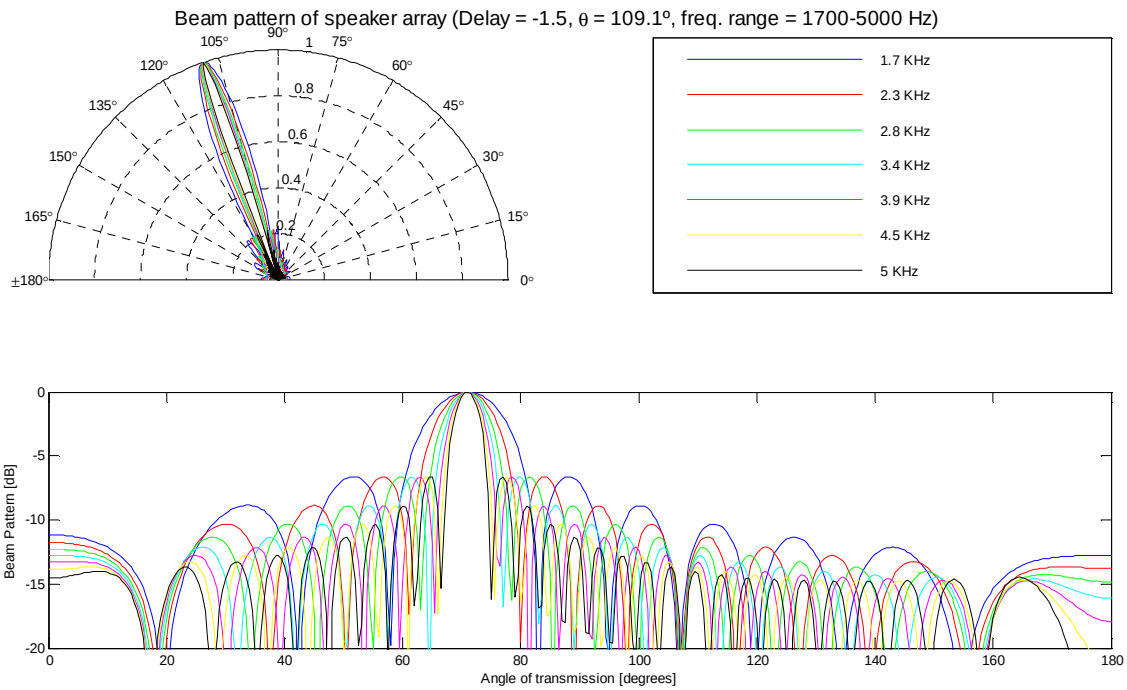

Beam pattern of speaker array (Delay = -4, θ = 150.7º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
1
0.8
0.6
0.4
0.2
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]

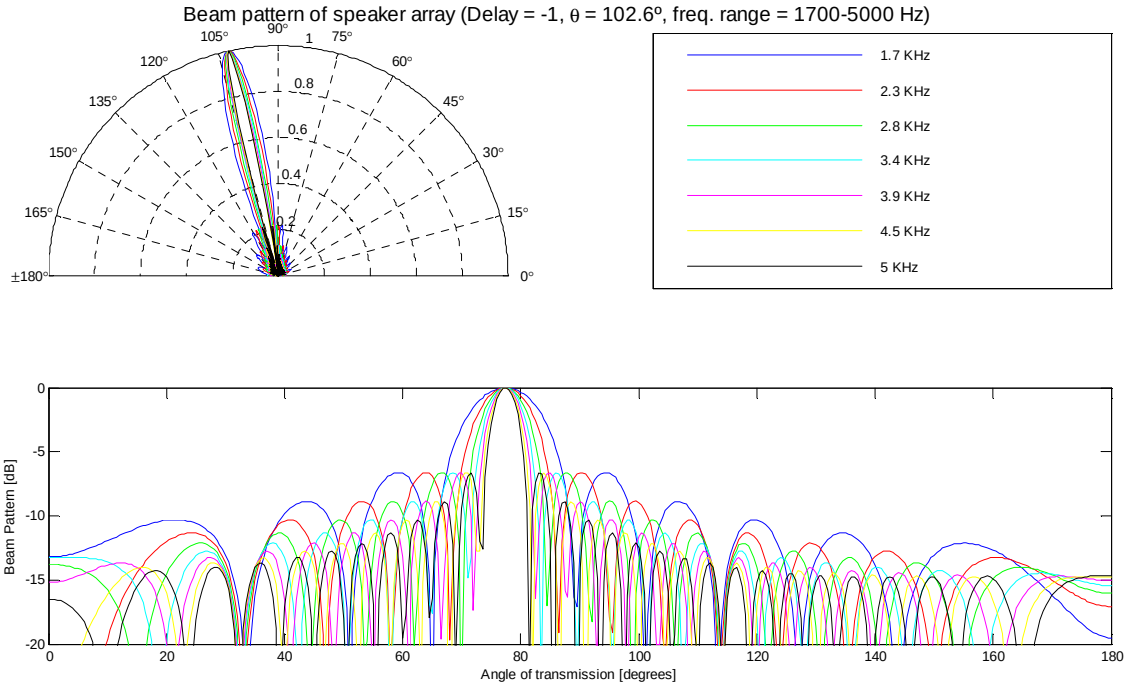

Beam pattern of speaker array (Delay = -3.5, θ = 139.7º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
1
0.8
0.6
0.4
0.2
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]

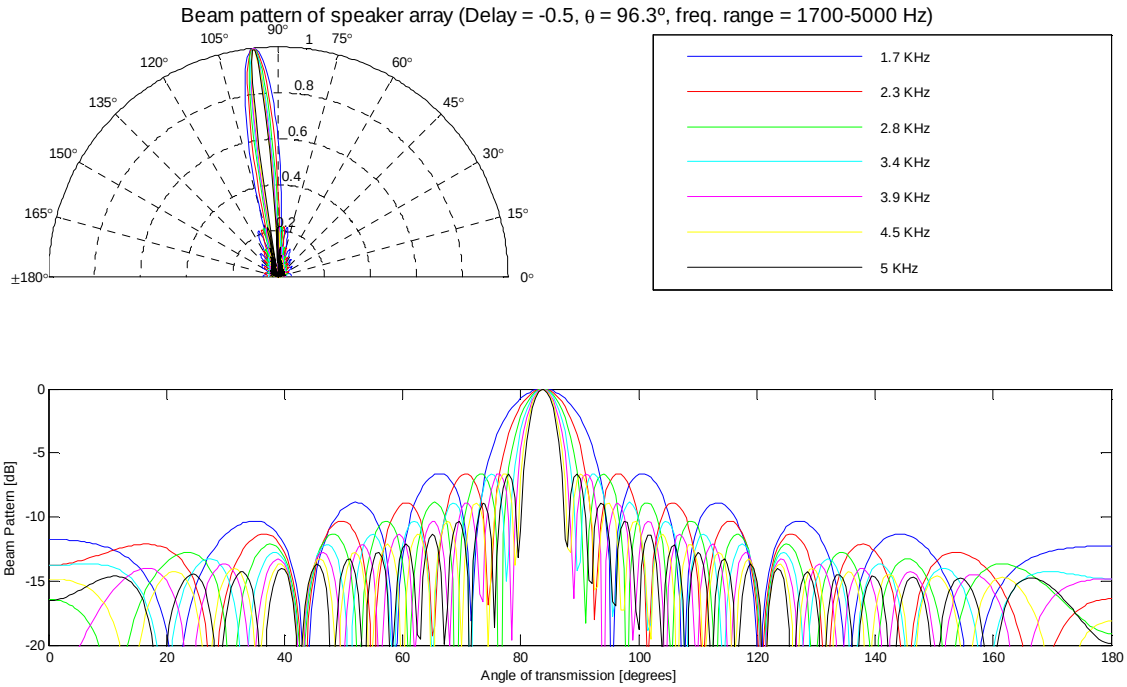

Beam pattern of speaker array (Delay = -3, θ = 130.8º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
1
0.8
0.6
0.4
0.2
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]


Beam pattern of speaker array (Delay = -2.5, θ = 123º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
1
0.8
0.6
0.4
0.2
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]


Beam pattern of speaker array (Delay = -2, θ = 115.8º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
0.8
0.6
0.4
0.2
1
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]


Beam pattern of speaker array (Delay = -1.5, θ = 109.1º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
0.8
0.6
0.4
0.2
1
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]


Beam pattern of speaker array (Delay = -1, θ = 102.6º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
1
0.8
0.6
0.4
0.2
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]


Beam pattern of speaker array (Delay = -0.5, θ = 96.3º, freq. range = 1700-5000 Hz)
90°
105°
75°
120°
60°
135°
45°
150°
30°
165°
15°
±180°
0°
1
0.8
0.6
0.4
0.2
1.7 KHz
2.3 KHz
2.8 KHz
3.4 KHz
3.9 KHz
4.5 KHz
5 KHz
Beam Pattern [dB]
0
-5
-10
-15
-20
0
20
40
60
80
100
120
140
160
180
Angle of transmission [degrees]

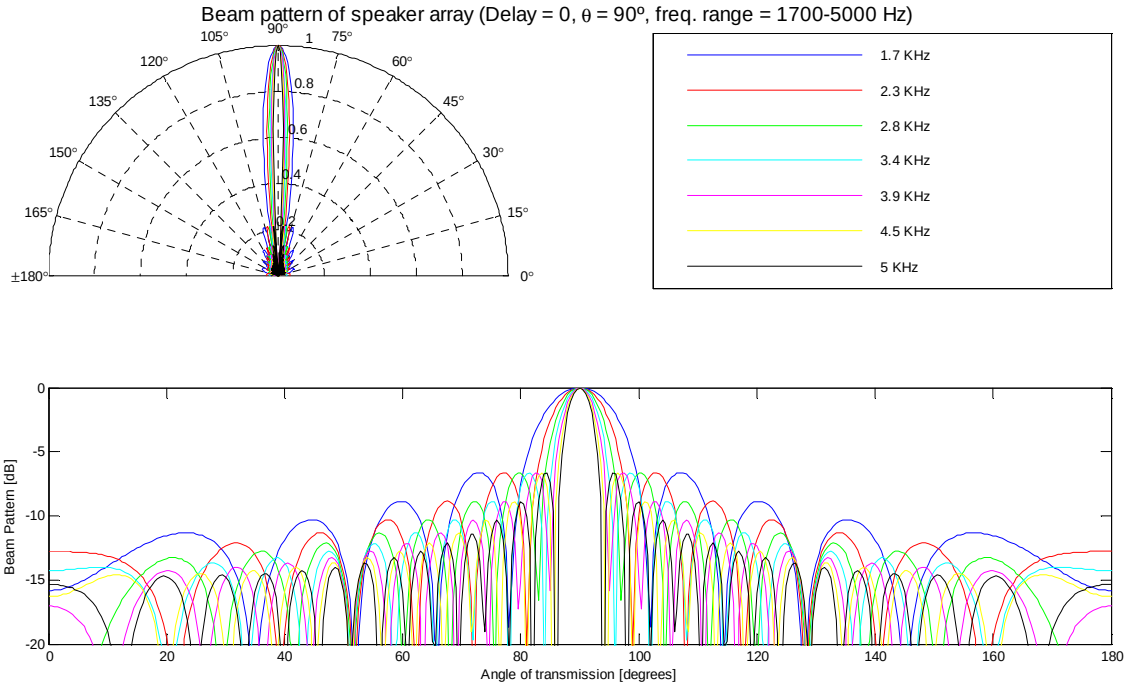



**Figure A.1.** Summary of the transmitted radiation patterns during the sounding phase. Due to symmetry, only the first ten directions of transmission are shown.

# Appendix B

# Manual

In order to initialize this GUI, MATLAB must be run and the "Current Directory" must be located in the same path as the GUI files. Then, the user has to write the following line in the prompt of MATLAB:

```
> main
```

An example of the window that appears is shown in Figure B.1. The dialog is divided in three sections: "Files", "Show results" and "Create signal". The "Exit" button located at the bottom closes the GUI and the communication with the DSP. The "Options" menu on the top left corner includes two fields:

- **Open.** It loads the signals sent and recorded on the microphone constellation in a previous session in order to analyze the results.
- **Exit.** It has the same effect as the "Exit" button.

## B.1 Files

This section is devoted to manage the communication between the GUI and the DSP. It loads, runs and halts the programs on the DSP. It also manages the process of loading the samples from the DSP to the PC in order to analyze them in MATLAB.

The files panel is divided two parts:

- **Sounding phase.** It runs Sounding_phase_48kHz.pjt on the DSP. The microphone constellation must be connected to the SAM-30.
    - **Load samples.** It loads the data from the DSP to MATLAB and stores them in a file. This option slows down Sounding_phase_48kHz.pjt while the data is loading, but it makes it possible to analyze it in MATLAB. Otherwise, the

results cannot be obtained. This option does not affect the behavior of the system, as it is only needed to see the results graphically.

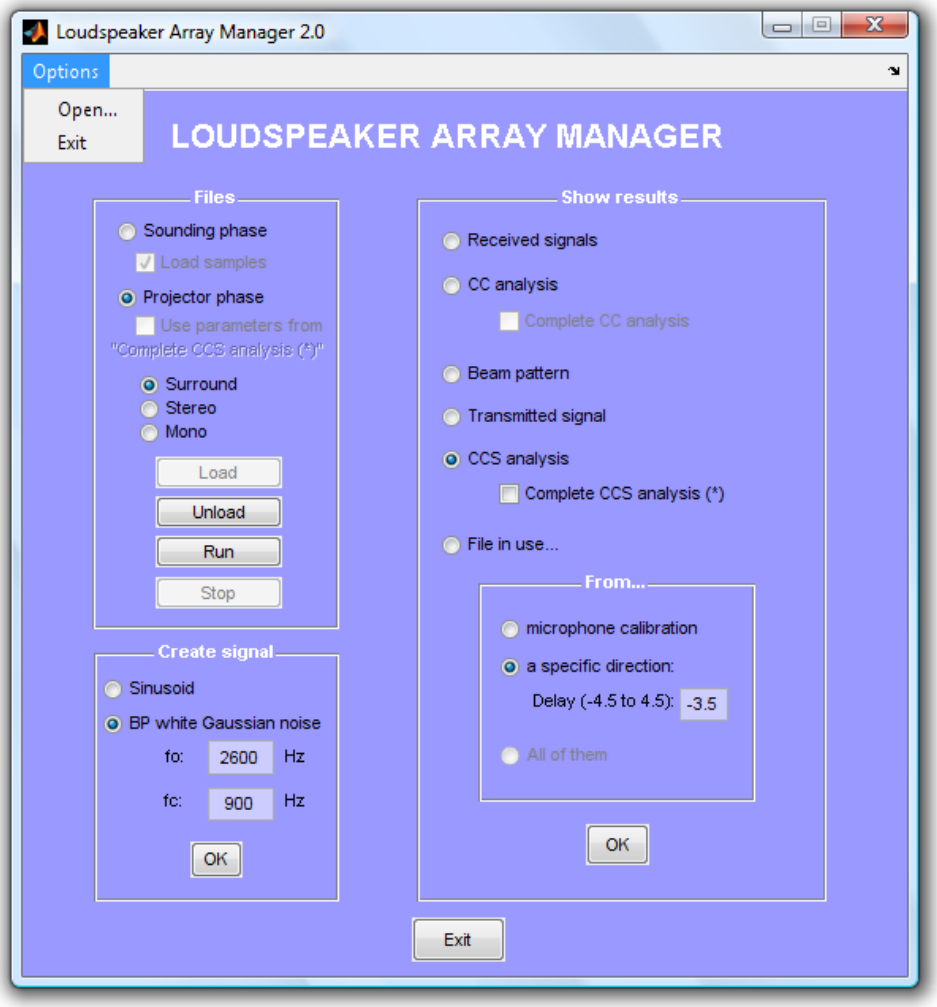



**Figure B.1.** Example of the GUI.

- **Projector phase.** It runs Projector_phase_48kHz.pjt on the DSP. The DVD-player must be connected to the SAM-30.
    - **Use parameters from "Complete CSS analysis (*)".** After running the complete analysis (in the "Show results" section) in MATLAB, this option is enabled. It permits to choose the `steering_delay`, `attenuation` and `transmit_delay` instead of the ones obtained from the DSP.

  There are three possible modes to play the sound:
    - **Surround.** It performs a 4.0-channel setup with the obtained parameters.
    - **Stereo.** It performs a 2.0-channel setup with the obtained parameters.
    - **Mono.** It performs a 1.0-channel setup.

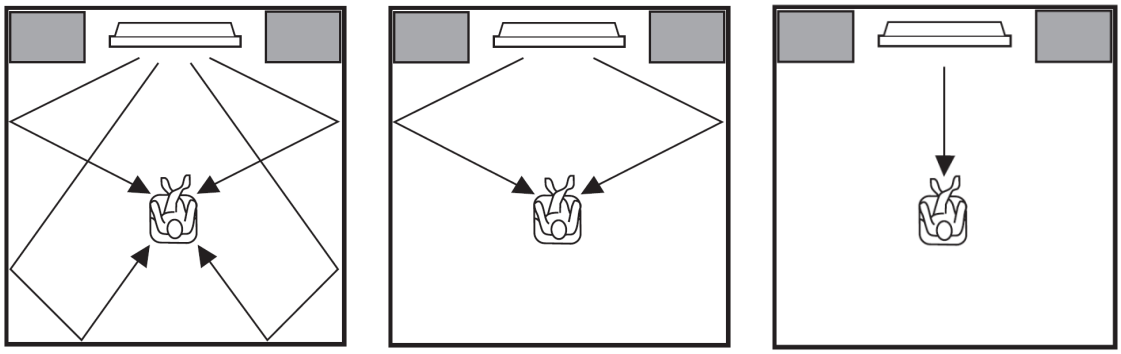


**Figure B.2.** From left to right: surround, stereo and mono setup.

To manage these options, four buttons are available:

- **Load.** It loads the selected program (Sounding_phase_48kHz.pjt or Projector_phase_48kHz.pjt) in the DSP.
- **Unload.** It unloads the current program from the DSP.
- **Run.** It runs the program loaded in the DSP.
- **Stop.** It stops the execution of the program in the DSP.

## B.2 Show Results

This dialog can only be used when the samples are loaded on MATLAB. There are two ways to do it: as explained above, by clicking in the "Load samples" checkbox before running the study of the room, or in the "Options" menu by loading a file stored in previous sessions.

When the samples are loaded in MATLAB the following options are enabled:

- **Received signals.** It shows the four received signals at the matrix of microphones in the time and frequency domains.
- **CC analysis.** It performs the TDOA method explained in section 3.3.1 for the specified delay at the sound projector. It shows the radiation pattern of the sound recorded at the matrix of microphones.
    - **Complete CC analysis.** It carries out the complete analysis done in the DSP. It illustrates the sound power received in some specific angles for each direction of transmission.
- **Beam pattern.** It illustrates the theoretical Array Factor (AF) of the sound projector for the frequency range of the sent signal and for the specified direction of transmission at the loudspeaker array.
- **Transmitted signal.** It shows the sent signal during the sounding phase in the time and frequency domains.
- **CSS analysis.** It performs the TDOA method explained in section 3.3.3 for the specified delay at the sound projector. It shows the radiation pattern of the sound recorded at the matrix of microphones and the propagation delay-DOA graphic.

    - **Complete CSS analysis.** It carries out a complete analysis of this TDOA method. It illustrates the sound power received in some specific angles for each direction of transmission. Moreover, it obtains the `attenuation`, `steering_delay` and `transmit_delay` variables calculated for this method. It enables the "Use parameters from Complete CSS analysis (*)" checkbox in the "Files" section.
- **File in use.** It shows the path and the name of the file where the samples are stored on the PC.

Depending on the selected option, the signal to be analyzed can be chosen in the "From…" dialog located inside the "Show results" dialog. There are three options:

- **Microphone calibration.** It corresponds to the samples recorded during first transmission of the signal, when the microphones are calibrated.
- **A specific direction.** It corresponds to the samples recorded during the delay specified below this option.
    - **Delay (-4.5 to 4.5).** It indicates the delay in samples between the speakers applied at the sound projector during the sounding phase. The possible values include the interval ranged from -4.5 to 4.5 with steps of 0.5.
- **All of them.** This option is only available when "Received signals" is selected.

## B.3 Create Signal

This dialog permits the user to choose the kind of signal to transmit for the study of the room. By default the signal presented in subsection 4.1 is selected. The options are:

- **Sinusoid.** Only $f_o$ can be modified. It creates a sinusoid with $f_o$ as a central frequency.
- **BP (Band-Pass) white Gaussian noise.** Both $f_o$ and $f_c$ can be modified. It creates a Band-Pass WGN centered in $f_o$ and with cut-off frequencies at $f_{c1} = f_o - f_c$ and $f_{c2} = f_o + f_c$.

# Appendix C

# Additional Tables

The total amount of values obtained from the experiments performed in a quiet and noisy environments are shown in function of the technique used (CC or CSS techniques).

As explained in section Chapter 6, the channels 3 and 4 obtained bad results for this laboratory room and they are discarded. The blue tables show all the values of the variables (including channels 3 and 4), and the red tables only show the values which apply to channels 1 and 2.

## C.1 Additional Tables Obtained from CC Technique

### C.1.1 Experiments Performed in a Quiet Environment

### C.1.1.1 Values Corresponding to `steering_delay` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 1.1** | -3.5 | -2 | 2.5 | 2.5 |
| **Number 1.2** | -3.5 | -2 | 3 | 3 |
| **Number 1.3** | -3.5 | -2 | 2.5 | 2.5 |
| **Number 1.4** | -3.5 | -2 | 4.5 | 4.5 |
| **Number 1.5** | -3.5 | -2 | 4 | 4 |
| **Number 1.6** | -3.5 | -2 | 4 | 4 |
| **Number 1.7** | -3.5 | -2 | 4 | 4 |
| **Number 1.8** | -3.5 | -2 | 4 | 4 |
| **Number 1.9** | -3.5 | -2 | 4 | 4 |
| **Number 1.10** | -3.5 | -2 | 4 | 4 |

**Table C.1.** Summary of the values corresponding to `steering_delay` obtained from the ten experiments in a quiet environment using the CC technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 1.1** | -3.5 | -2 |
| **Number 1.2** | -3.5 | -2 |
| **Number 1.3** | -3.5 | -2 |
| **Number 1.4** | -3.5 | -2 |
| **Number 1.5** | -3.5 | -2 |
| **Number 1.6** | -3.5 | -2 |
| **Number 1.7** | -3.5 | -2 |
| **Number 1.8** | -3.5 | -2 |
| **Number 1.9** | -3.5 | -2 |
| **Number 1.10** | -3.5 | -2 |

**Table C.2.** Summary of the values corresponding to `steering_delay` obtained from the ten experiments in a quiet environment using the CC technique. Only channels 1 and 2 are shown.

### C.1.1.2 Values Corresponding to `attenuation` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 1.1** | 0.4293 | 0.6151 | 1.0000 | 1.0000 |
| **Number 1.2** | 0.3482 | 0.5614 | 1.0000 | 1.0000 |
| **Number 1.3** | 0.4000 | 0.5795 | 1.0000 | 1.0000 |
| **Number 1.4** | 0.2309 | 0.3740 | 1.0000 | 1.0000 |
| **Number 1.5** | 0.3459 | 0.5287 | 1.0000 | 1.0000 |
| **Number 1.6** | 0.3563 | 0.5420 | 1.0000 | 1.0000 |
| **Number 1.7** | 0.3683 | 0.5374 | 1.0000 | 1.0000 |
| **Number 1.8** | 0.3447 | 0.5312 | 1.0000 | 1.0000 |
| **Number 1.9** | 0.3501 | 0.5727 | 1.0000 | 1.0000 |
| **Number 1.10** | 0.3826 | 0.5649 | 1.0000 | 1.0000 |

**Table C.3.** Summary of the values corresponding to `attenuation` obtained from the ten experiments in a quiet environment using the CC technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 1.1** | 0.6979 | 1.0000 |
| **Number 1.2** | 0.6202 | 1.0000 |
| **Number 1.3** | 0.6902 | 1.0000 |
| **Number 1.4** | 0.6173 | 1.0000 |
| **Number 1.5** | 0.6542 | 1.0000 |
| **Number 1.6** | 0.6573 | 1.0000 |
| **Number 1.7** | 0.6853 | 1.0000 |
| **Number 1.8** | 0.6489 | 1.0000 |
| **Number 1.9** | 0.6113 | 1.0000 |
| **Number 1.10** | 0.6772 | 1.0000 |
| **AVERAGE** | **0.6560** | **1.0000** |

**Table C.4.** Summary of the values corresponding to `attenuation` obtained from the ten experiments in a quiet environment using the CC technique. Only channels 1 and 2 are shown.

### C.1.2 Tests Performed in a Noisy Environment

Due to the CC technique does not work properly under noisy conditions, the values obtained during the five experiments are erroneous. Therefore, in this case, the red tables (the ones that only show channels 1 and 2) are not shown because they lack interest.

#### C.1.2.1 Values Corresponding to `steering_delay` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 2.1** | 0 | 0 | 4.5 | 4.5 |
| **Number 2.2** | 0 | 0 | 4 | 4.5 |
| **Number 2.3** | 0 | 0 | 4.5 | 4.5 |
| **Number 2.4** | 0 | 0 | 4.5 | 4.5 |
| **Number 2.5** | 0 | 0 | 4.5 | 4.5 |

**Table C.5.** Summary of the values corresponding to `steering_delay` obtained from the five experiments in a noisy environment using the CC technique.

#### C.1.2.2 Values Corresponding to `attenuation` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 2.1** | 0.7402 | 0.7402 | 1.0000 | 1.0000 |
| **Number 2.2** | 0.7032 | 0.7032 | 0.9872 | 1.0000 |
| **Number 2.3** | 0.7697 | 0.7697 | 1.0000 | 1.0000 |
| **Number 2.4** | 0.7281 | 0.7281 | 1.0000 | 1.0000 |
| **Number 2.5** | 0.7392 | 0.7392 | 1.0000 | 1.0000 |
| **AVERAGE** | **0.7361** | **0.7361** | **0.9974** | **1.0000** |

**Table C.6.** Summary of the values corresponding to `attenuation` obtained from the five experiments in a noisy environment using the CC technique.

## C.2 Additional Tables Obtained from CSS Technique

### C.2.1 Tests Performed in a Quiet Environment

#### C.2.1.1 Values Corresponding to `steering_delay` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 1.1** | -3.5 | -2 | 2.5 | 4 |
| **Number 1.2** | -3.5 | -2 | 4.5 | 4.5 |
| **Number 1.3** | -3.5 | -2 | 4 | 4 |
| **Number 1.4** | -3.5 | -2 | 4.5 | 4.5 |
| **Number 1.5** | -3.5 | -2 | 4 | 4 |
| **Number 1.6** | -3.5 | -2 | 4 | 4 |
| **Number 1.7** | -3.5 | -2 | 4 | 4 |
| **Number 1.8** | -3.5 | -2 | 2 | 2 |
| **Number 1.9** | -3.5 | -2 | 4 | 4 |
| **Number 1.10** | -3.5 | -2 | 4 | 4 |

**Table C.7.** Summary of the values corresponding to `steering_delay` obtained from the ten experiments in a quiet environment using the CSS technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 1.1** | -3.5 | -2 |
| **Number 1.2** | -3.5 | -2 |
| **Number 1.3** | -3.5 | -2 |
| **Number 1.4** | -3.5 | -2 |
| **Number 1.5** | -3.5 | -2 |
| **Number 1.6** | -3.5 | -2 |
| **Number 1.7** | -3.5 | -2 |
| **Number 1.8** | -3.5 | -2 |
| **Number 1.9** | -3.5 | -2 |
| **Number 1.10** | -3.5 | -2 |

**Table C.8.** Summary of the values corresponding to `steering_delay` obtained from the ten experiments in a quiet environment using the CSS technique. Only channels 1 and 2 are shown.

### C.2.1.2 Values Corresponding to `attenuation` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 1.1** | 0.3015 | 0.4427 | 1.0000 | 1.0000 |
| **Number 1.2** | 0.3140 | 0.4777 | 1.0000 | 1.0000 |
| **Number 1.3** | 0.3214 | 0.4830 | 1.0000 | 1.0000 |
| **Number 1.4** | 0.3212 | 0.5073 | 1.0000 | 1.0000 |
| **Number 1.5** | 0.3203 | 0.5201 | 1.0000 | 1.0000 |
| **Number 1.6** | 0.3169 | 0.5541 | 1.0000 | 1.0000 |
| **Number 1.7** | 0.3149 | 0.5657 | 1.0000 | 1.0000 |
| **Number 1.8** | 0.3831 | 0.6805 | 1.0000 | 1.0000 |
| **Number 1.9** | 0.3140 | 0.5570 | 1.0000 | 1.0000 |
| **Number 1.10** | 0.3135 | 0.5796 | 1.0000 | 1.0000 |

**Table C.9.** Summary of the values corresponding to `attenuation` obtained from the ten experiments in a quiet environment using the CSS technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 1.1** | 0.6810 | 1.0000 |
| **Number 1.2** | 0.6573 | 1.0000 |
| **Number 1.3** | 0.6654 | 1.0000 |
| **Number 1.4** | 0.6331 | 1.0000 |
| **Number 1.5** | 0.6158 | 1.0000 |
| **Number 1.6** | 0.6546 | 1.0000 |
| **Number 1.7** | 0.6619 | 1.0000 |
| **Number 1.8** | 0.6379 | 1.0000 |
| **Number 1.9** | 0.6447 | 1.0000 |
| **Number 1.10** | 0.6403 | 1.0000 |
| **AVERAGE** | **0.6492** | **1.0000** |

**Table C.10.** Summary of the values corresponding to `attenuation` obtained from the ten experiments in a quiet environment using the CSS technique. Only channels 1 and 2 are shown.

### C.2.1.3 Values Corresponding to `transmit_delay` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 1.1** | 386 | 0 | 7 | 184 |
| **Number 1.2** | 1766 | 1384 | 0 | 0 |
| **Number 1.3** | 383 | 0 | 183 | 183 |
| **Number 1.4** | 1774 | 1389 | 0 | 0 |
| **Number 1.5** | 390 | 0 | 185 | 185 |
| **Number 1.6** | 405 | 0 | 207 | 207 |
| **Number 1.7** | 1766 | 1362 | 0 | 0 |
| **Number 1.8** | 383 | 0 | 394 | 394 |
| **Number 1.9** | 384 | 0 | 213 | 213 |
| **Number 1.10** | 403 | 0 | 232 | 232 |

**Table C.11.** Summary of the values corresponding to `transmit_delay` obtained from the ten experiments in a quiet environment using the CSS technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 1.1** | 386 | 0 |
| **Number 1.2** | 382 | 0 |
| **Number 1.3** | 383 | 0 |
| **Number 1.4** | 385 | 0 |
| **Number 1.5** | 390 | 0 |
| **Number 1.6** | 405 | 0 |
| **Number 1.7** | 404 | 0 |
| **Number 1.8** | 383 | 0 |
| **Number 1.9** | 384 | 0 |
| **Number 1.10** | 403 | 0 |
| **AVERAGE** | **391** | **0** |

**Table C.12.** Summary of the values corresponding to `transmit_delay` obtained from the ten experiments in a quiet environment using the CSS technique. Only channels 1 and 2 are shown.

## C.2.2 Tests Performed in a Noisy Environment

### C.2.2.1 Values Corresponding to `steering_delay` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 2.1** | -3.5 | -2 | 4.5 | 4 |
| **Number 2.2** | -3.5 | -2 | 4 | 4.5 |
| **Number 2.3** | -3.5 | -2 | 4 | 4 |
| **Number 2.4** | -3.5 | -2 | 4.5 | 4.5 |
| **Number 2.5** | -3.5 | -2 | 4 | 4 |

**Table C.13.** Summary of the values corresponding to `steering_delay` obtained from the five experiments in a noisy environment using the CSS technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 2.1** | -3.5 | -2 |
| **Number 2.2** | -3.5 | -2 |
| **Number 2.3** | -3.5 | -2 |
| **Number 2.4** | -3.5 | -2 |
| **Number 2.5** | -3.5 | -2 |

**Table C.14.** Summary of the values corresponding to `steerind_delay` obtained from the five experiments in a noisy environment using the CSS technique. Only channels 1 and 2 are shown.

### C.2.2.2 Values Corresponding to `attenuation` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 2.1** | 0.4320 | 0.6378 | 0.9815 | 1.0000 |
| **Number 2.2** | 0.4102 | 0.6291 | 0.9938 | 1.0000 |
| **Number 2.3** | 0.4098 | 0.6317 | 1.0000 | 1.0000 |
| **Number 2.4** | 0.4238 | 0.6482 | 1.0000 | 1.0000 |
| **Number 2.5** | 0.4190 | 0.6421 | 1.0000 | 1.0000 |

**Table C.15.** Summary of the values corresponding to `attenuation` obtained from the five experiments in a noisy environment using the CSS technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 1.1** | 0.6773 | 1.0000 |
| **Number 1.2** | 0.6520 | 1.0000 |
| **Number 1.3** | 0.6487 | 1.0000 |
| **Number 1.4** | 0.6538 | 1.0000 |
| **Number 1.5** | 0.6525 | 1.0000 |
| **AVERAGE** | **0.6568** | **1.0000** |

**Table C.16.** Summary of the values corresponding to `attenuation` obtained from the five experiments in a noisy environment using the CSS technique. Only channels 1 and 2 are shown.

### C.2.2.3 Values Corresponding to `transmit_delay` Variable

| Experiment | Channel 1 (front right) | Channel 2 (rear right) | Channel 3 (rear left) | Channel 4 (front left) |
|---|---|---|---|---|
| **Number 2.1** | 1591 | 1206 | 732 | 0 |
| **Number 2.2** | 1173 | 788 | 0 | 1909 |
| **Number 2.3** | 1290 | 900 | 0 | 0 |
| **Number 2.4** | 1425 | 1042 | 0 | 0 |
| **Number 2.5** | 1389 | 988 | 0 | 0 |

**Table C.17.** Summary of the values corresponding to `transmit_delay` obtained from the five experiments in a noisy environment using the CSS technique.

| Experiment | Channel 1 (front right) | Channel 2 (rear right) |
|---|---|---|
| **Number 1.1** | 385 | 0 |
| **Number 1.2** | 385 | 0 |
| **Number 1.3** | 390 | 0 |
| **Number 1.4** | 383 | 0 |
| **Number 1.5** | 401 | 0 |
| **AVERAGE** | **389** | **0** |

**Table C.18.** Summary of the values corresponding to `transmit_delay` obtained from the five experiments in a noisy environment using the CSS technique. Only channels 1 and 2 are shown.

# Appendix D

# Additional Graphs

This appendix includes the complete collection of plots obtained from two experiments performed in a quiet and noisy environment. These results were obtained during the sound tests performed in the laboratory described in section 2.1.

## D.1 Additional Graphs Obtained from CC Technique

### D.1.1 Tests Performed in a Quiet Environment

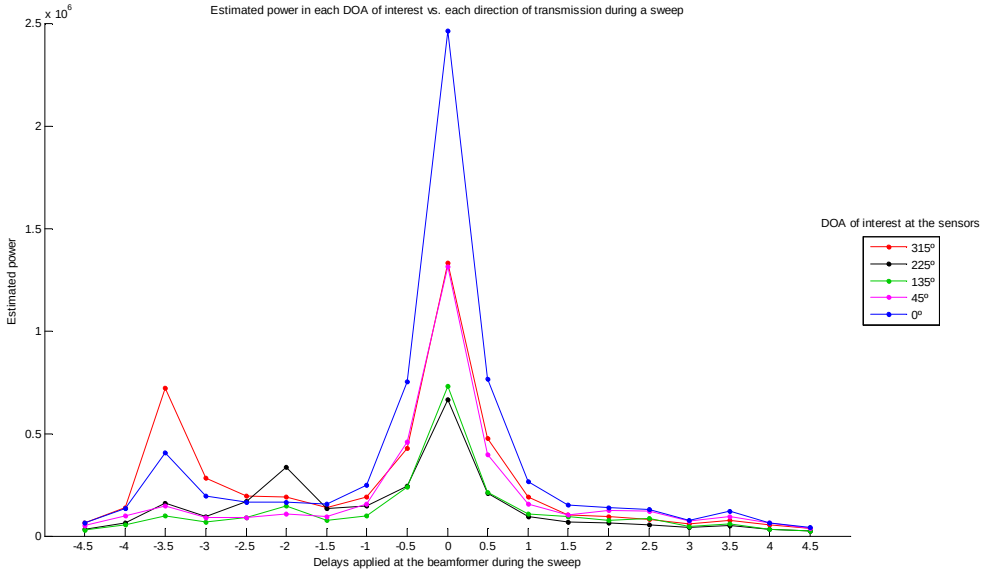



**Figure D.1.** Example of the estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) in a quiet environment. It has been obtained from the CC technique.

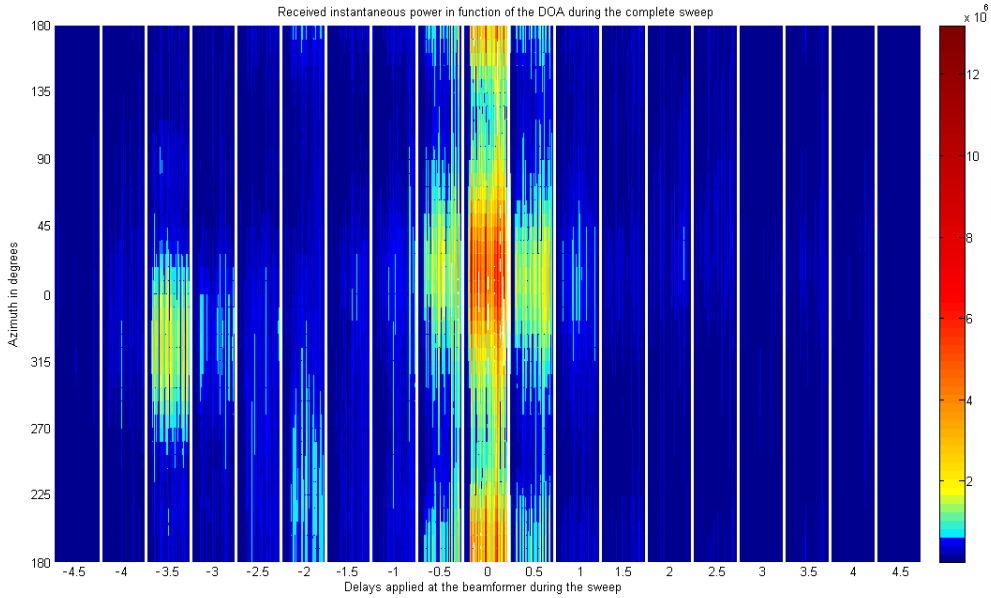
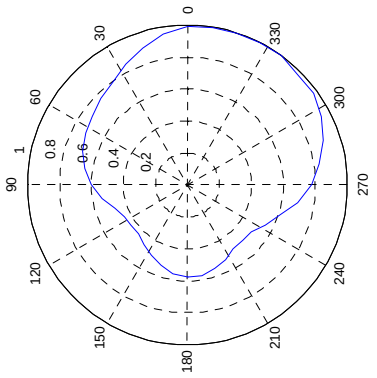
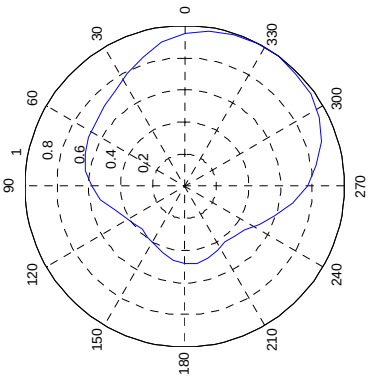


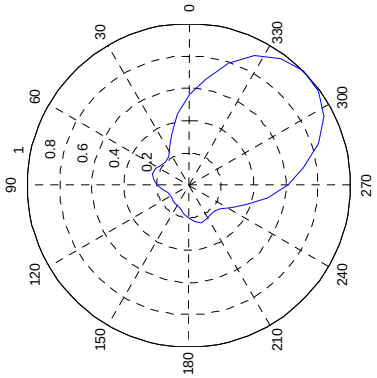

**Figure D.2.** Time-DOA plot of the received instantaneous sound power during the sounding phase in a quiet environment. It has been obtained from the CC technique.



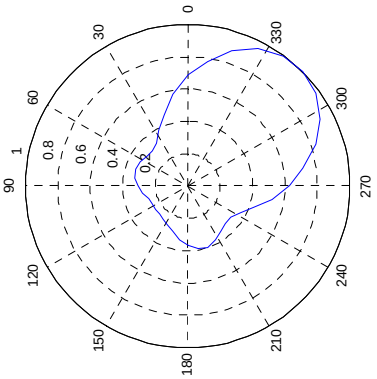
Steering delay = -4.5



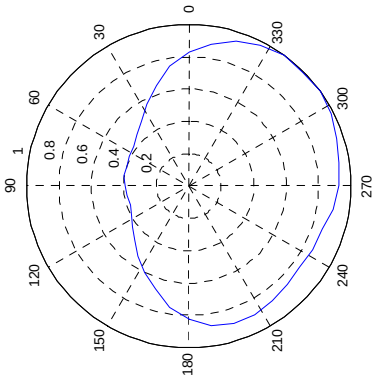
Steering delay = -4



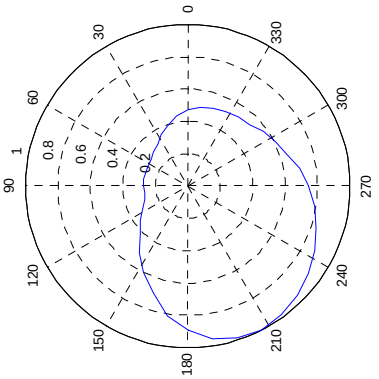
Steering delay = -3.5



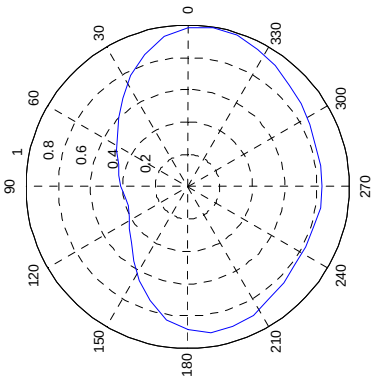
Steering delay = -3



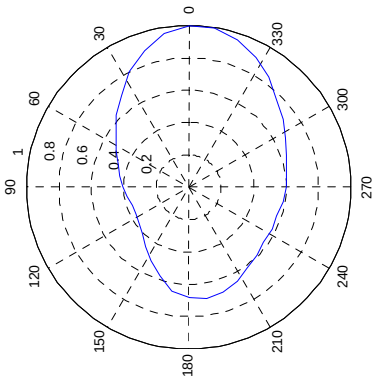
Steering delay = -2.5



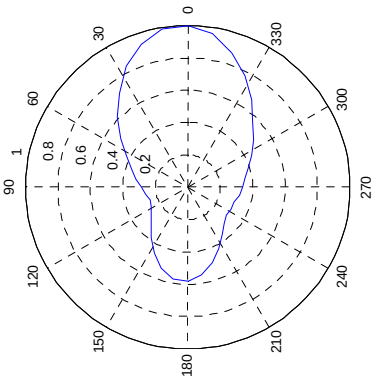
Steering delay = -2



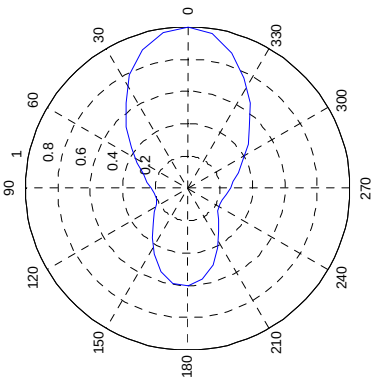
Steering delay = -1.5



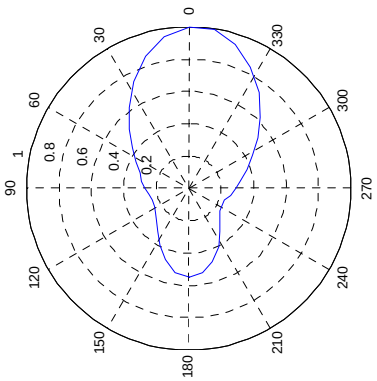
Steering delay = -1



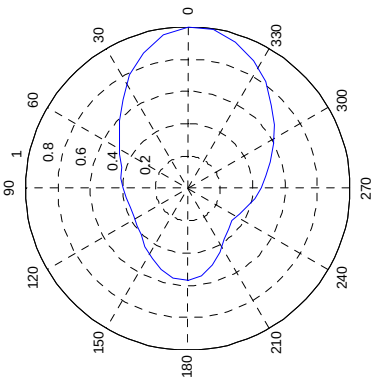
Steering delay = -0.5



Steering delay = 0



Steering delay = 0.5



Steering delay = 1

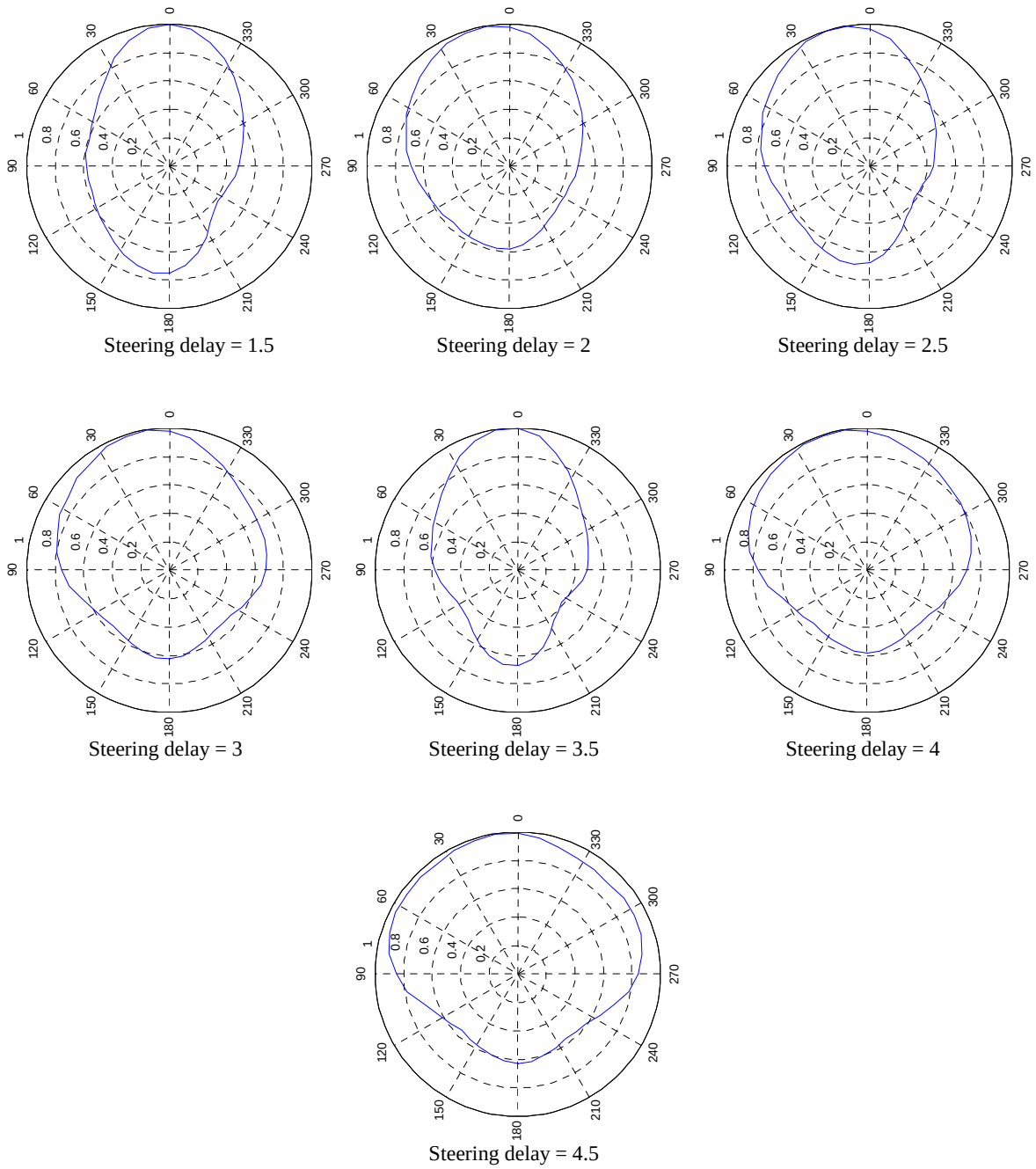



**Figure D.3.** Summary of the received radiation patterns during the sounding phase in a quiet environment using the CC technique.

## D.1.2 Tests Performed in a Noisy Environment

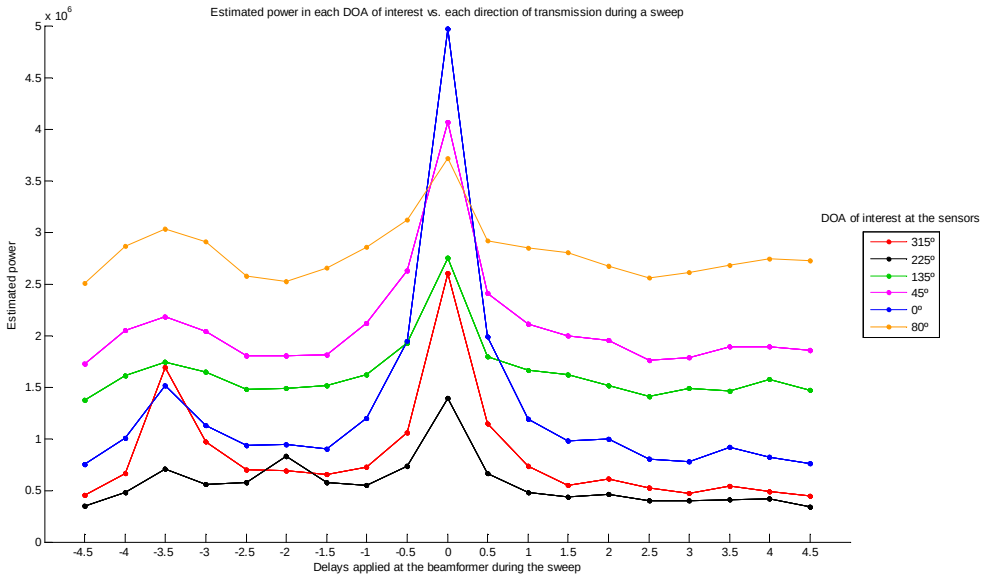



**Figure D.4.** Estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) in a noisy environment. It has been obtained from the CC technique.

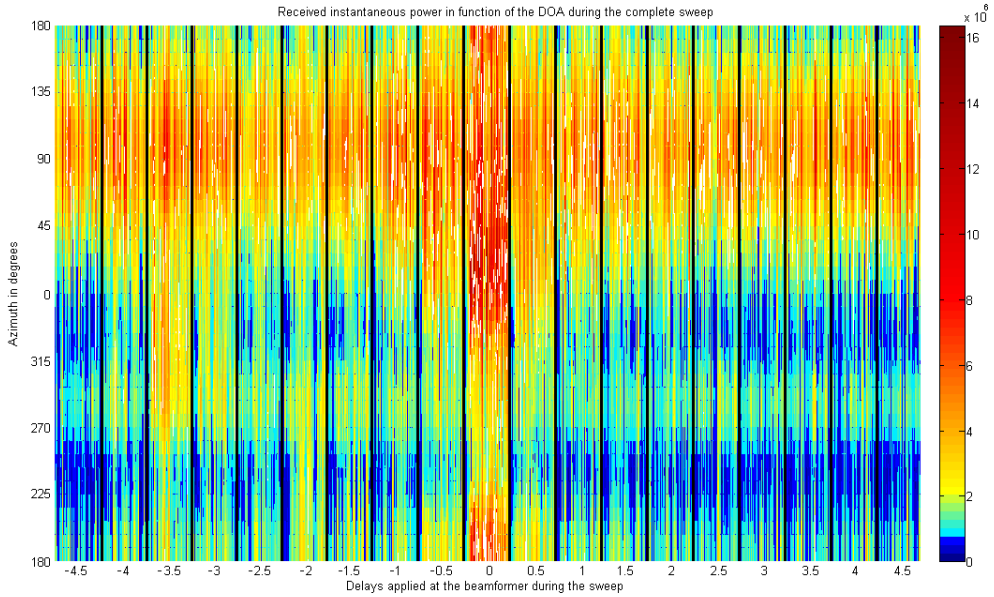
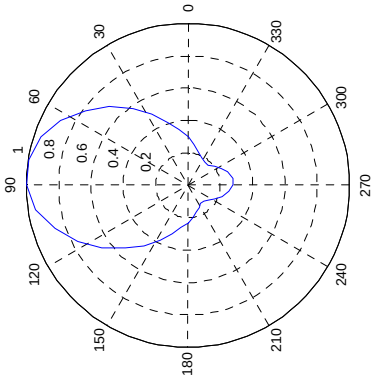
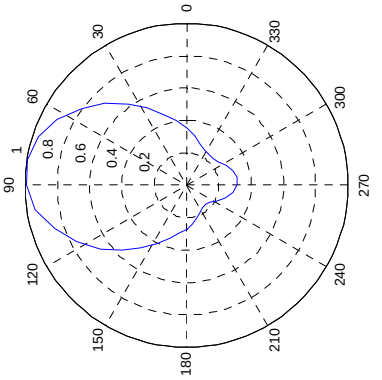


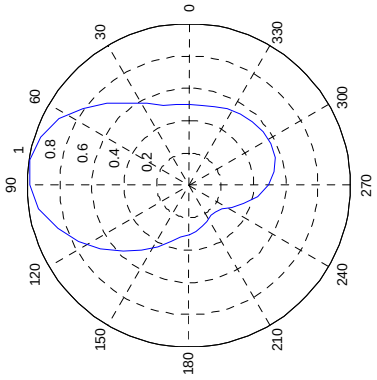

**Figure D.5.** Time-DOA plot of the received instantaneous sound power during the sounding phase in a noisy environment. It has been obtained from the CC technique.



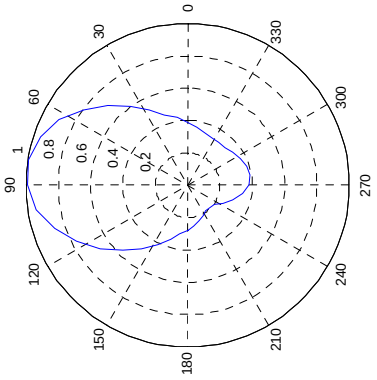
Steering delay = -4.5



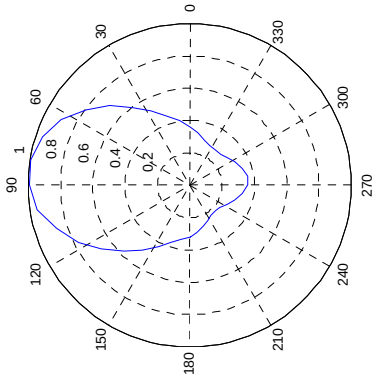
Steering delay = -4



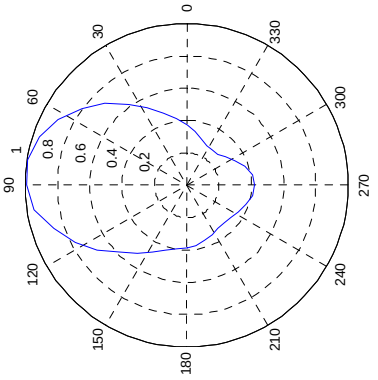
Steering delay = -3.5



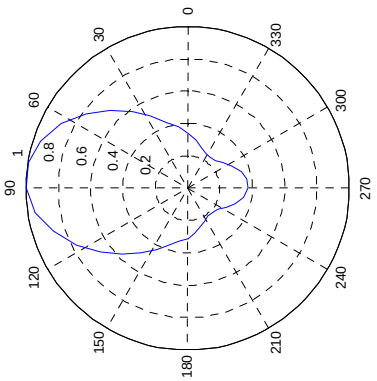
Steering delay = -3



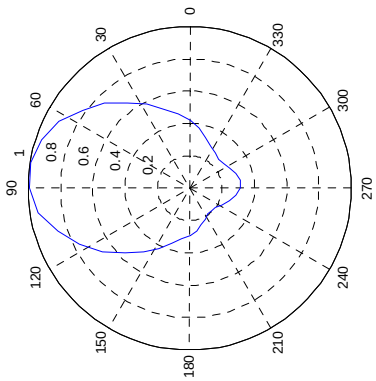
Steering delay = -2.5



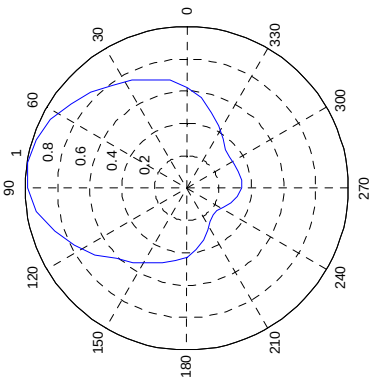
Steering delay = -2



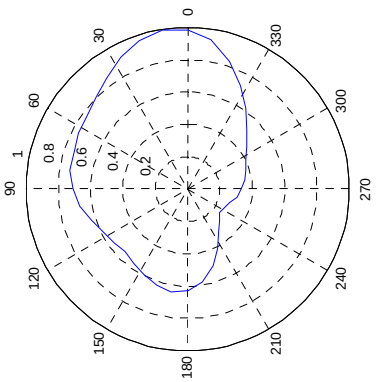
Steering delay = -1.5



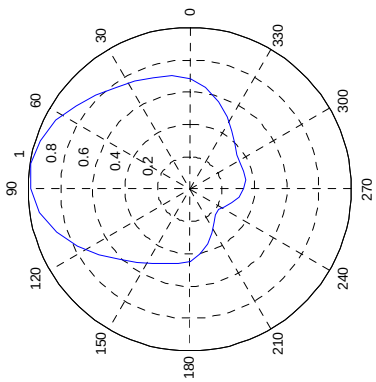
Steering delay = -1



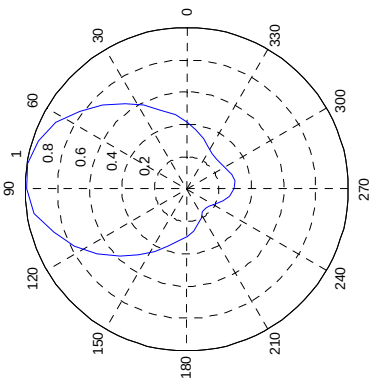
Steering delay = -0.5



Steering delay = 0



Steering delay = 0.5



Steering delay = 1

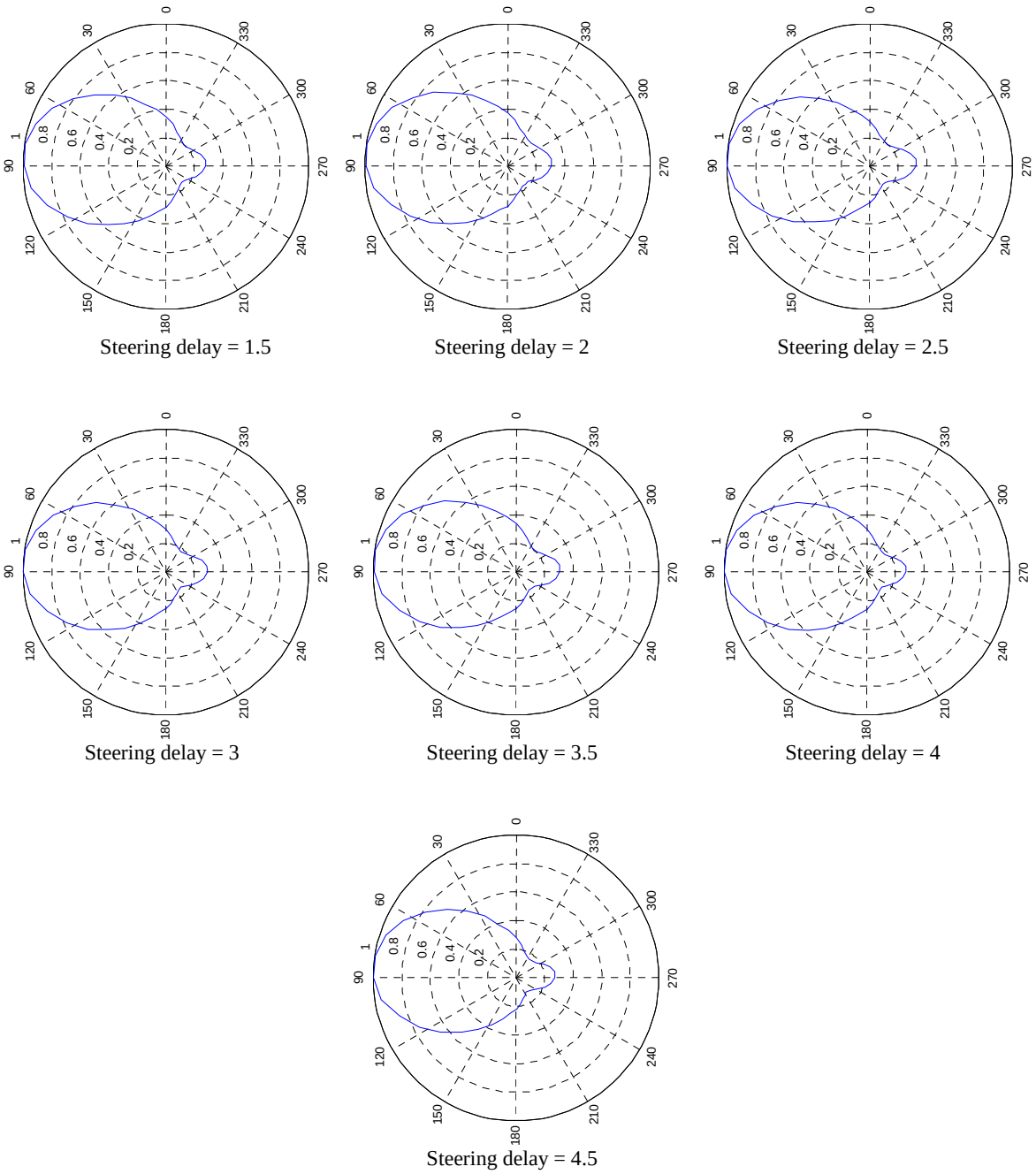



**Figure D.6.** Summary of the received radiation patterns during the sounding phase in a noisy environment using the CC technique.

## D.2 Additional Graphs Obtained from CSS Technique

### D.2.1 Tests Performed in a Quiet Environment

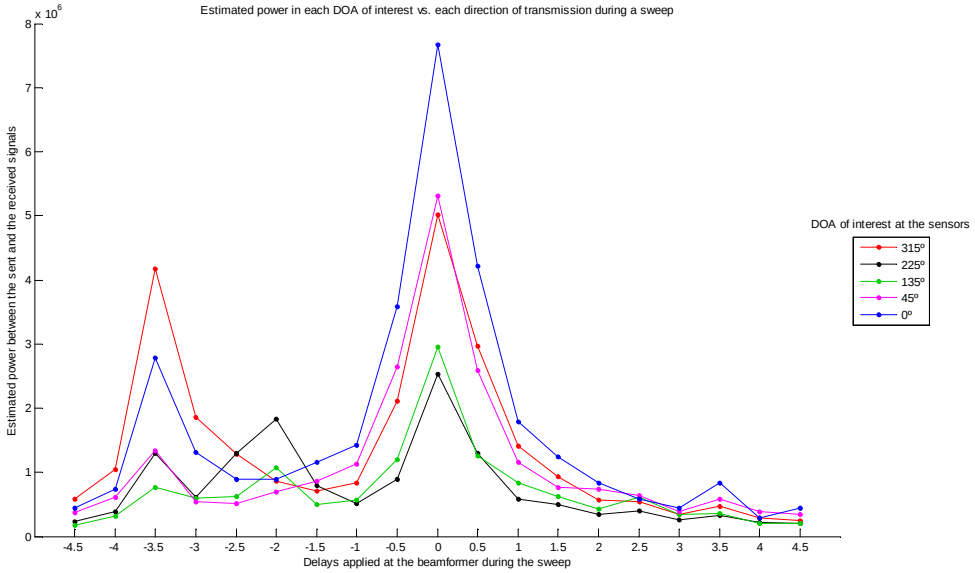
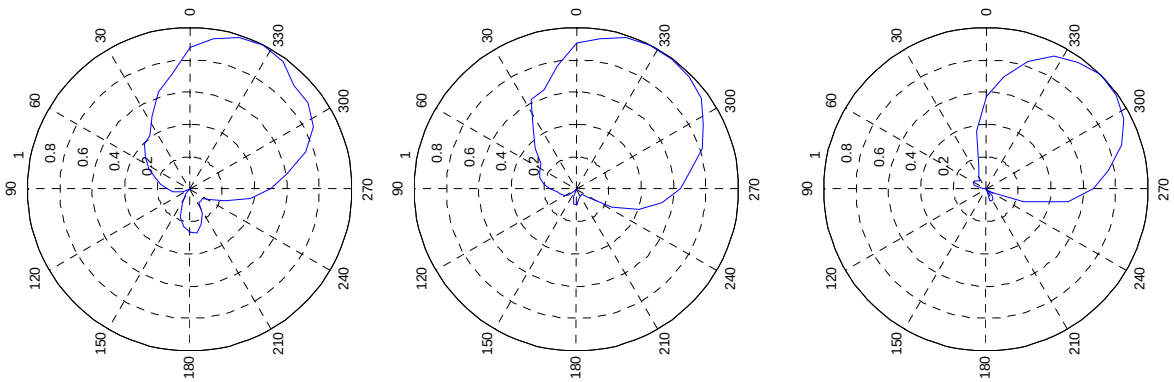
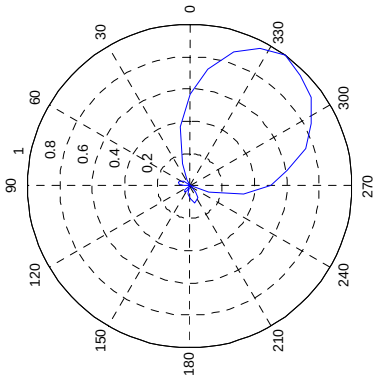


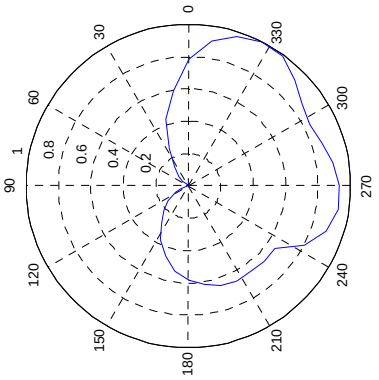

**Figure D.7.** Estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) for a quiet environment. It has been obtained from the CSS technique.



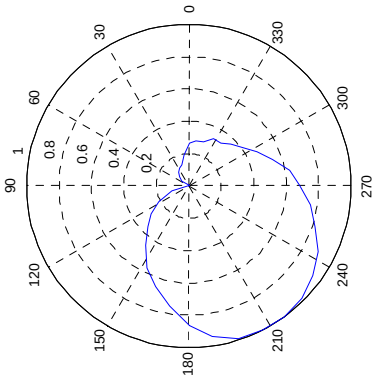
Steering delay = -4.5 Steering delay = -4 Steering delay = -3.5



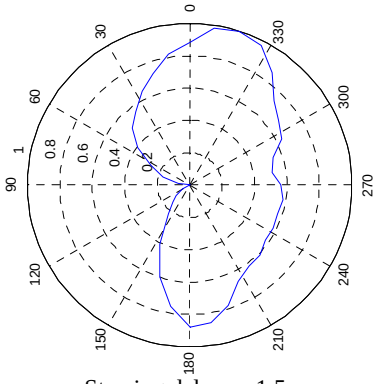
Steering delay = -3



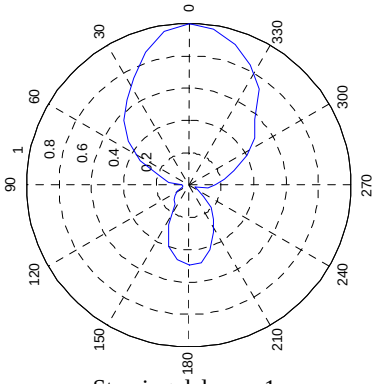
Steering delay = -2.5



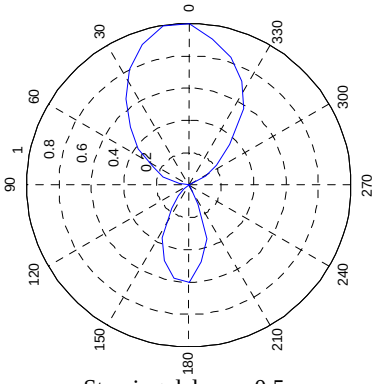
Steering delay = -2



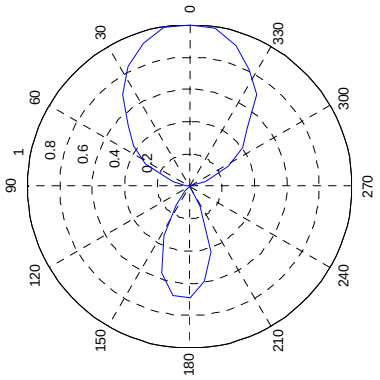
Steering delay = -1.5



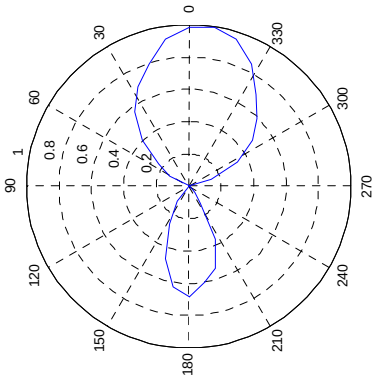
Steering delay = -1



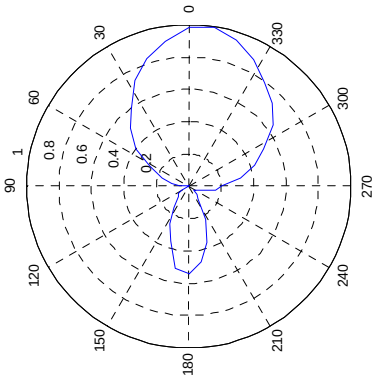
Steering delay = -0.5



Steering delay = 0



Steering delay = 0.5



Steering delay = 1

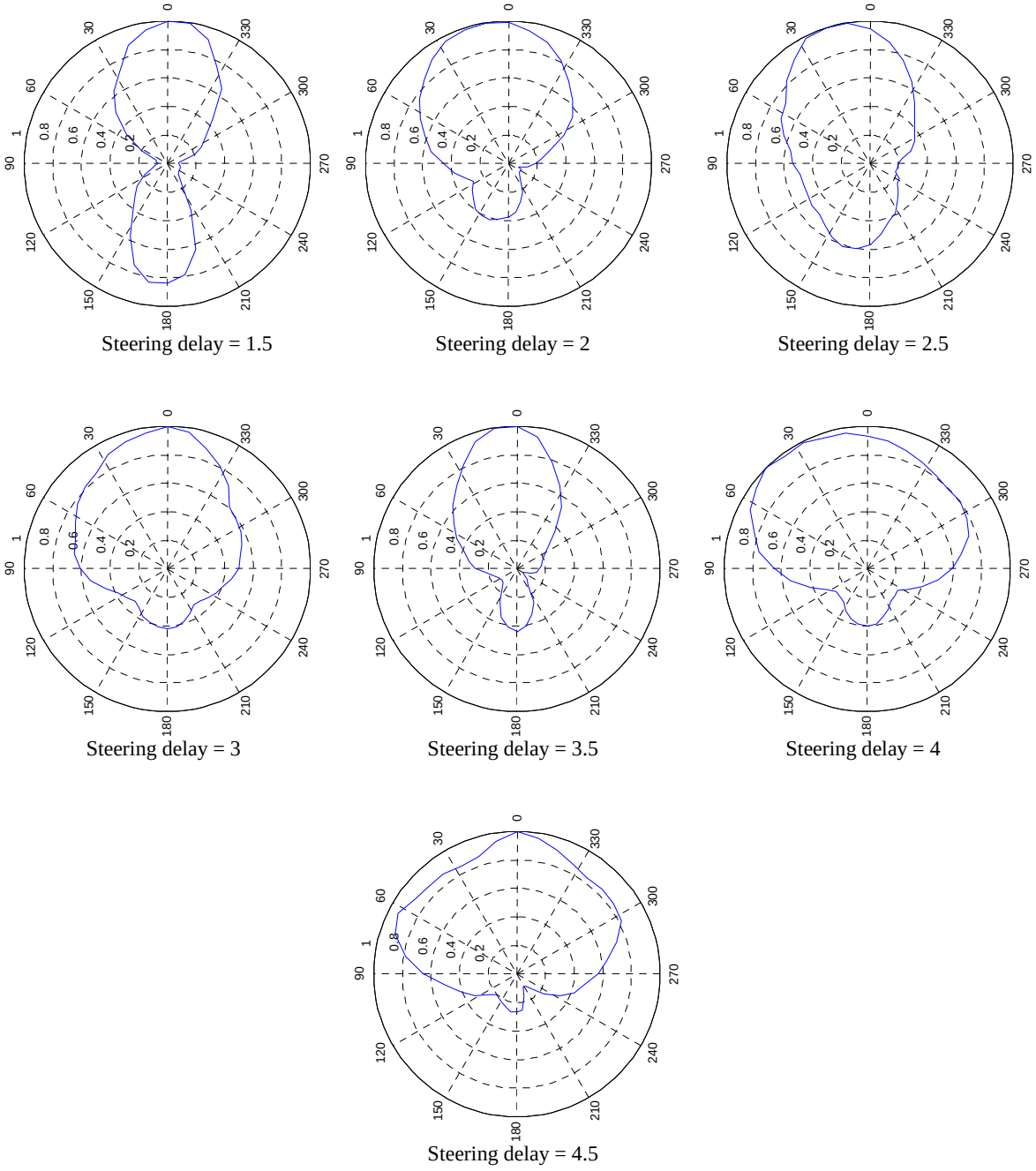
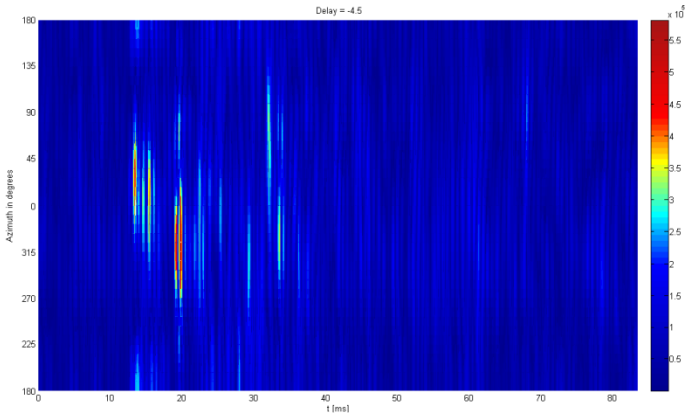
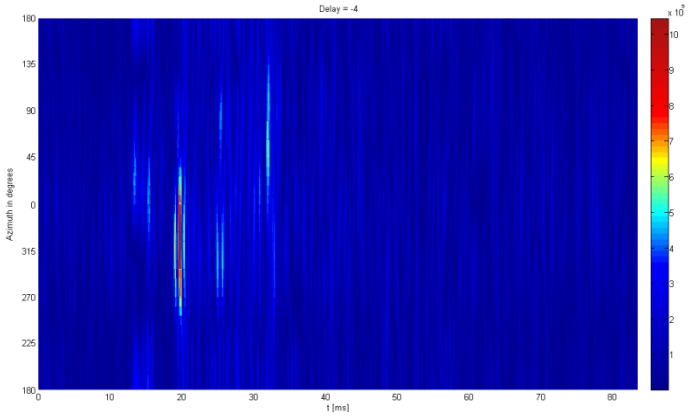



**Figure D.8.** Summary of the received radiation patterns during the sounding phase in a quiet environment using the CSS technique.

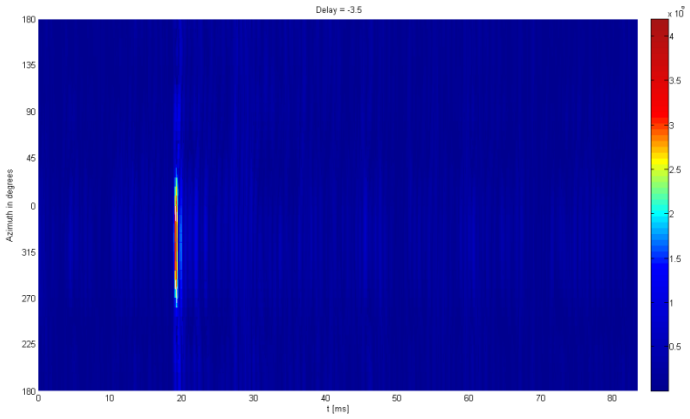
Steering delay = -4.5



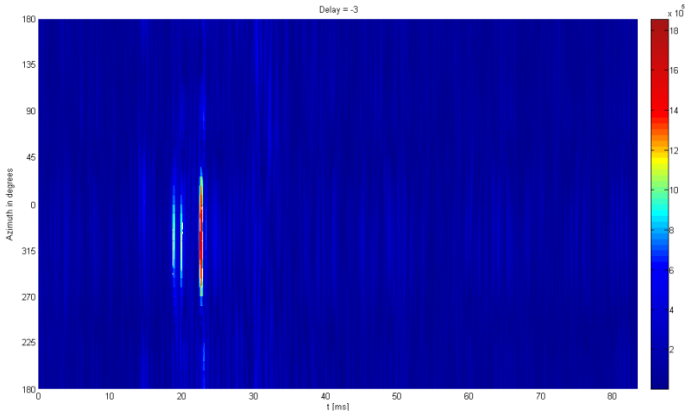

Steering delay = -4



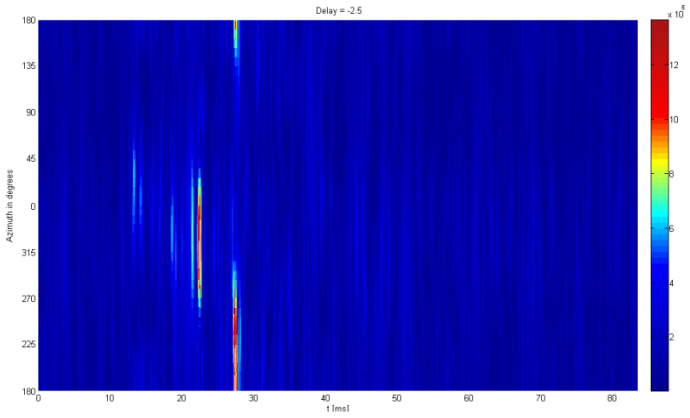

Steering delay = -3.5



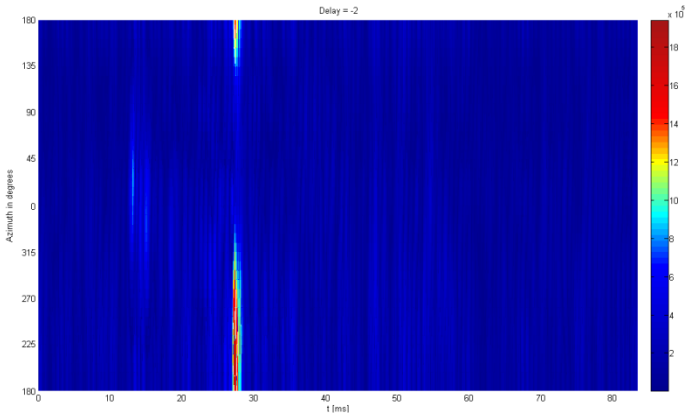

Steering delay = -3



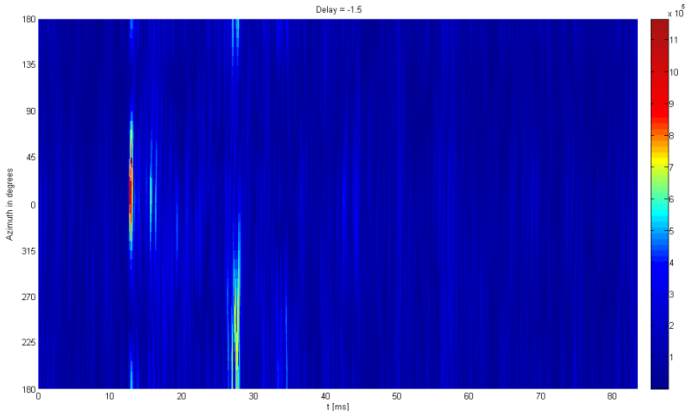
Steering delay = -2.5



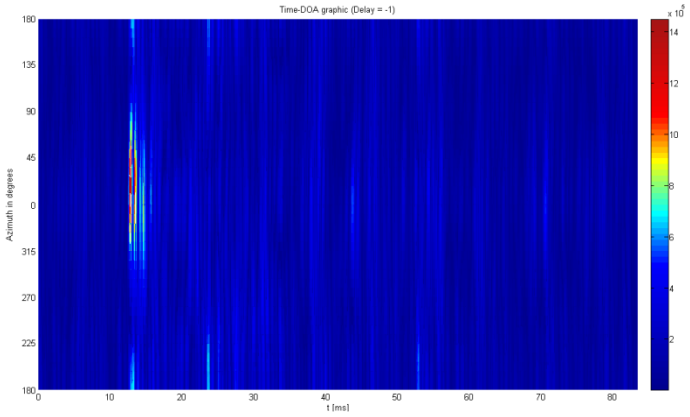

Steering delay = -2



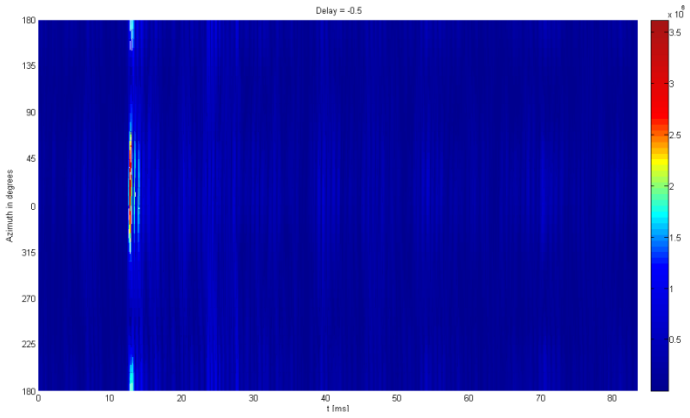

Steering delay = -1.5



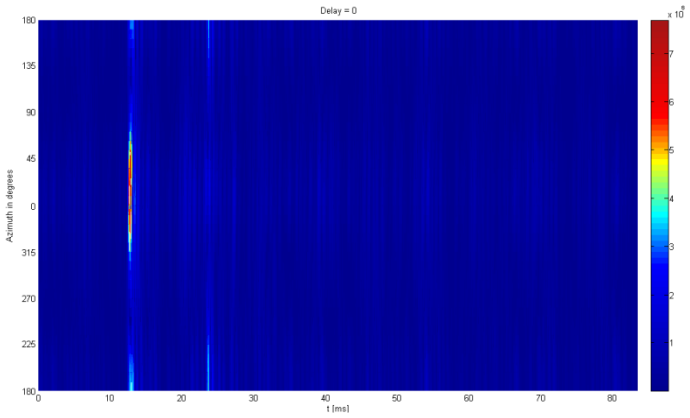

Steering delay = -1



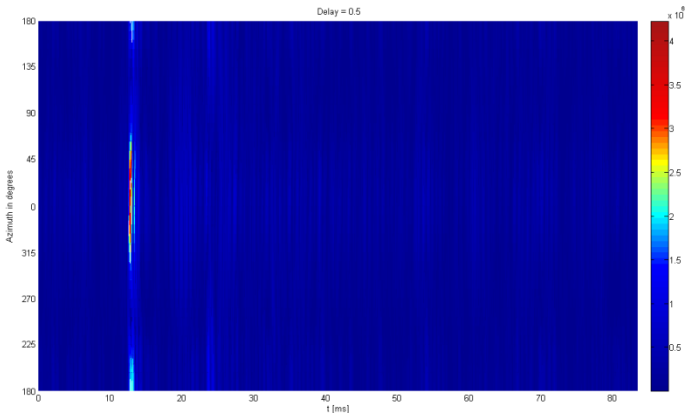
Steering delay = -0.5



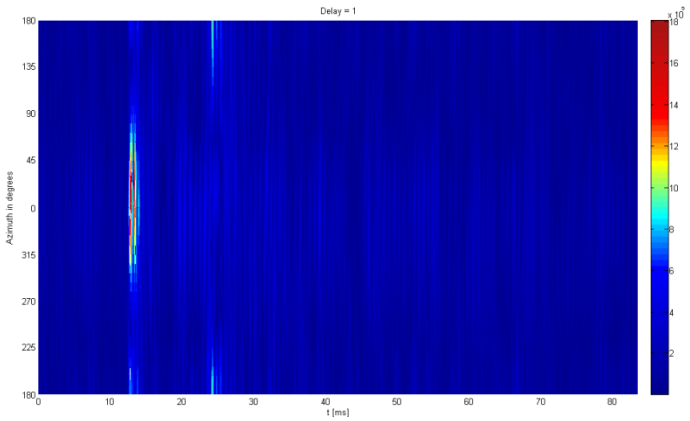

Steering delay = 0



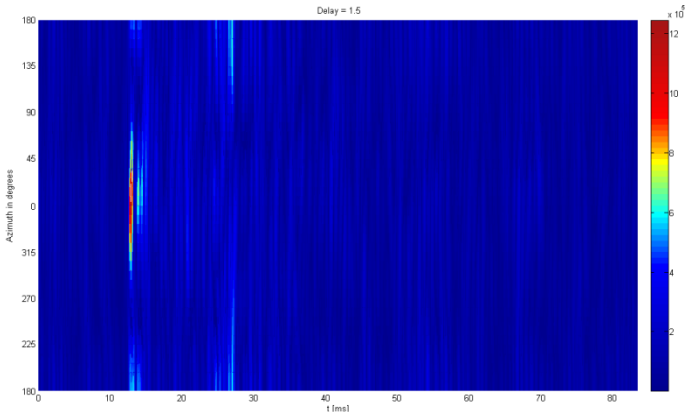

Steering delay = 0.5



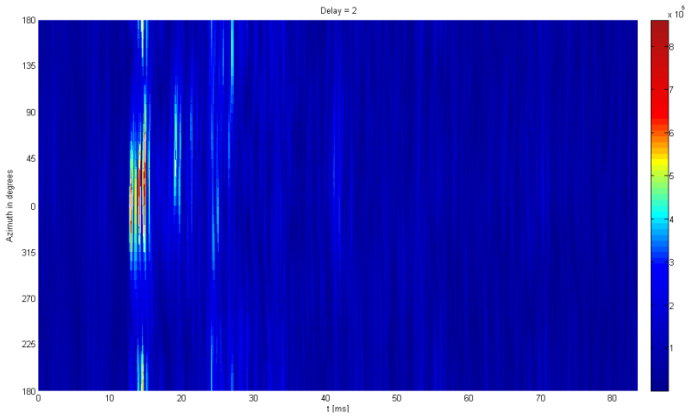

Steering delay = 1



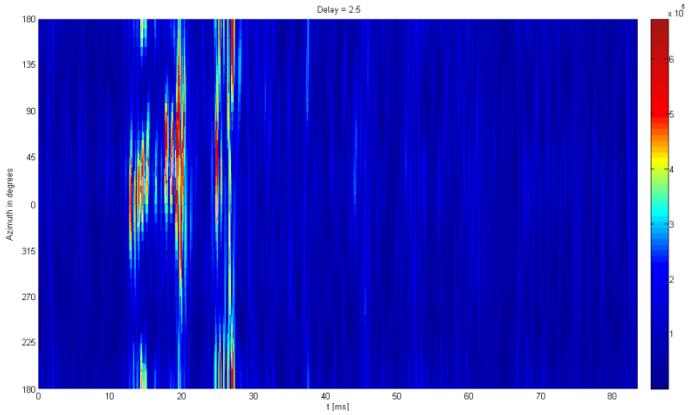
Steering delay = 1.5



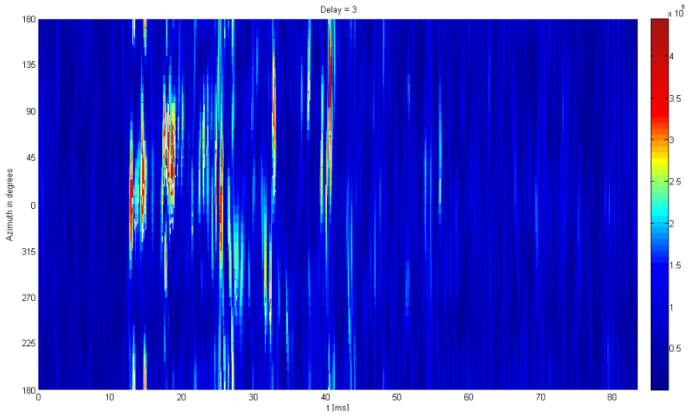

Steering delay = 2



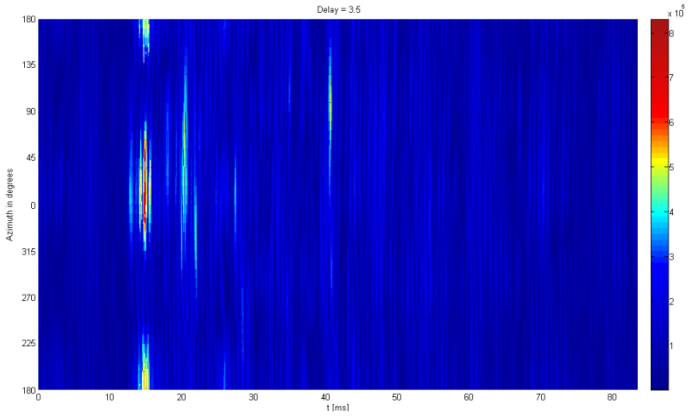

Steering delay = 2.5



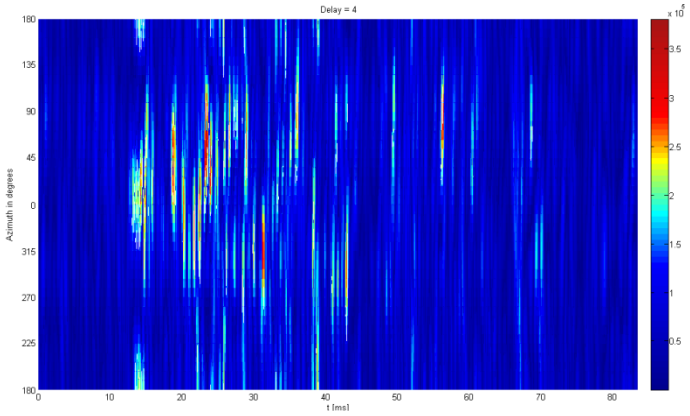

Steering delay = 3



Steering delay = 3.5




Steering delay = 4




Steering delay = 4.5

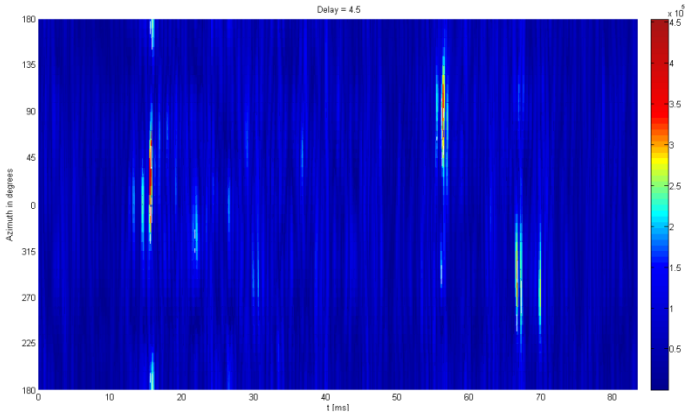



**Figure D.9.** Summary of the propagation delay-DOA plots received during the sounding phase in a quiet environment.

## D.2.2 Tests Performed in a Noisy Environment

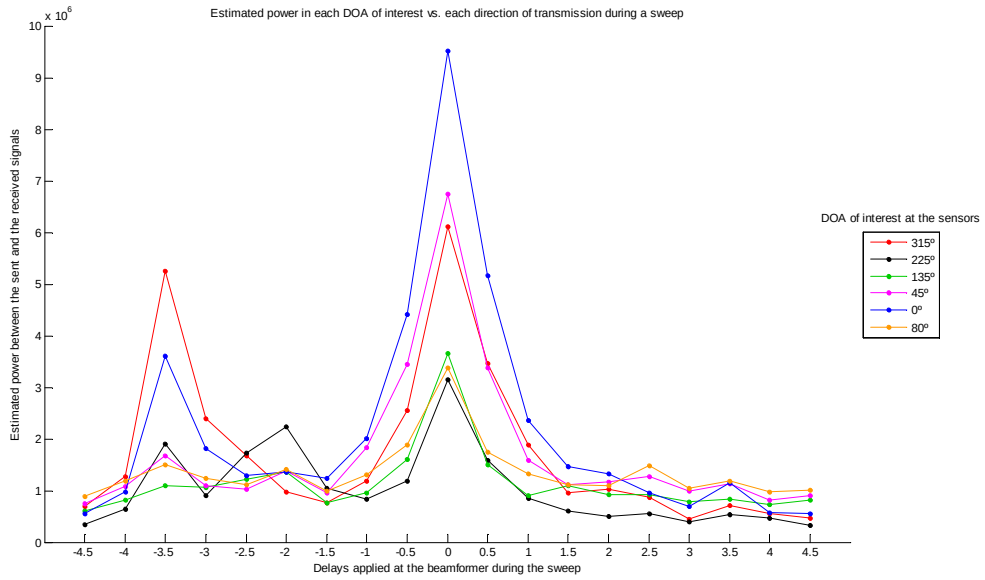
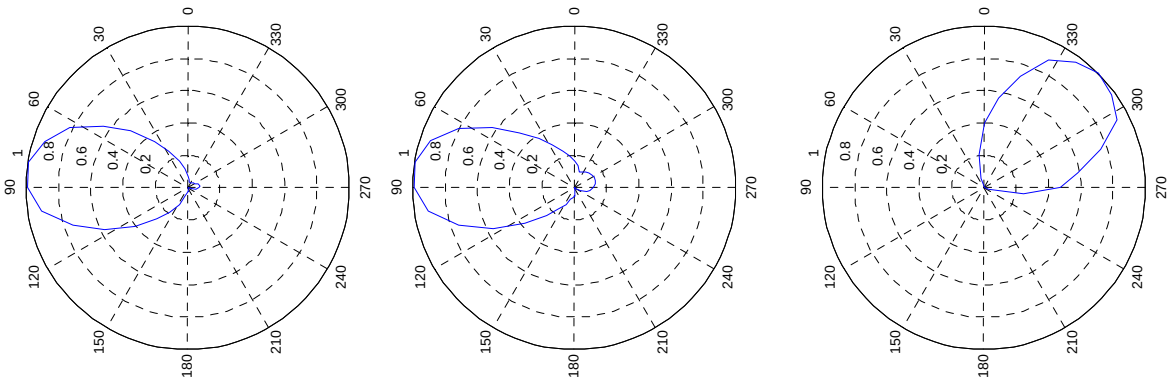
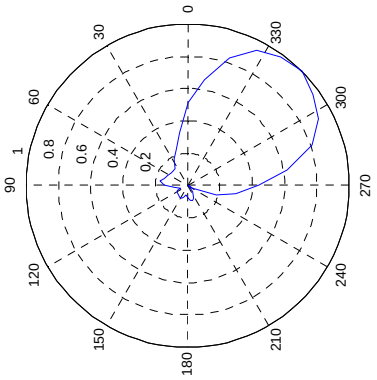


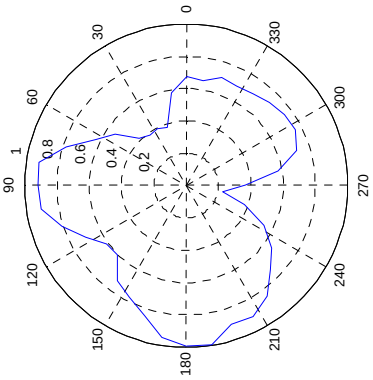

**Figure D.10.** Estimated power in each DOA of interest vs. each direction of transmission during the sounding phase (see Table 3.1) for a noisy environment. It has been obtained from the CSS technique.



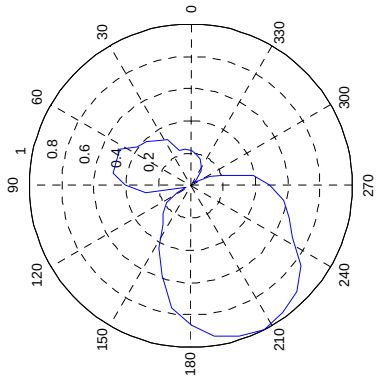
Steering delay = -4.5 Steering delay = -4 Steering delay = -3.5



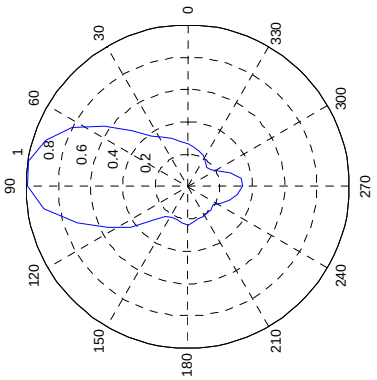

Steering delay = -3



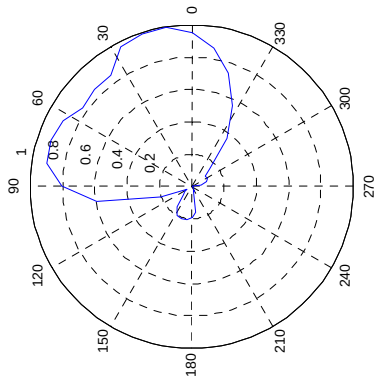

Steering delay = -2.5



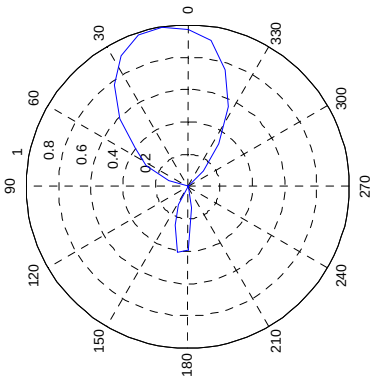

Steering delay = -2



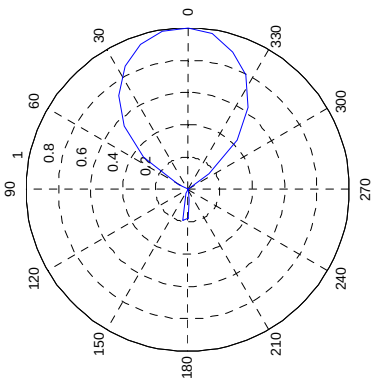

Steering delay = -1.5



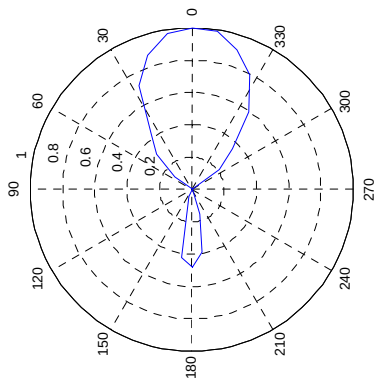

Steering delay = -1



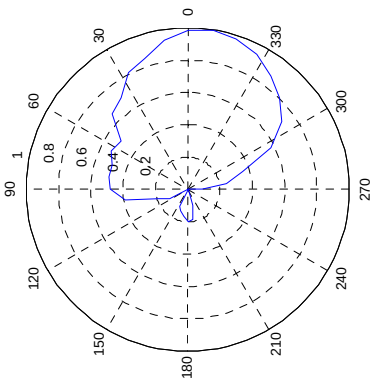

Steering delay = -0.5




Steering delay = 0




Steering delay = 0.5




Steering delay = 1

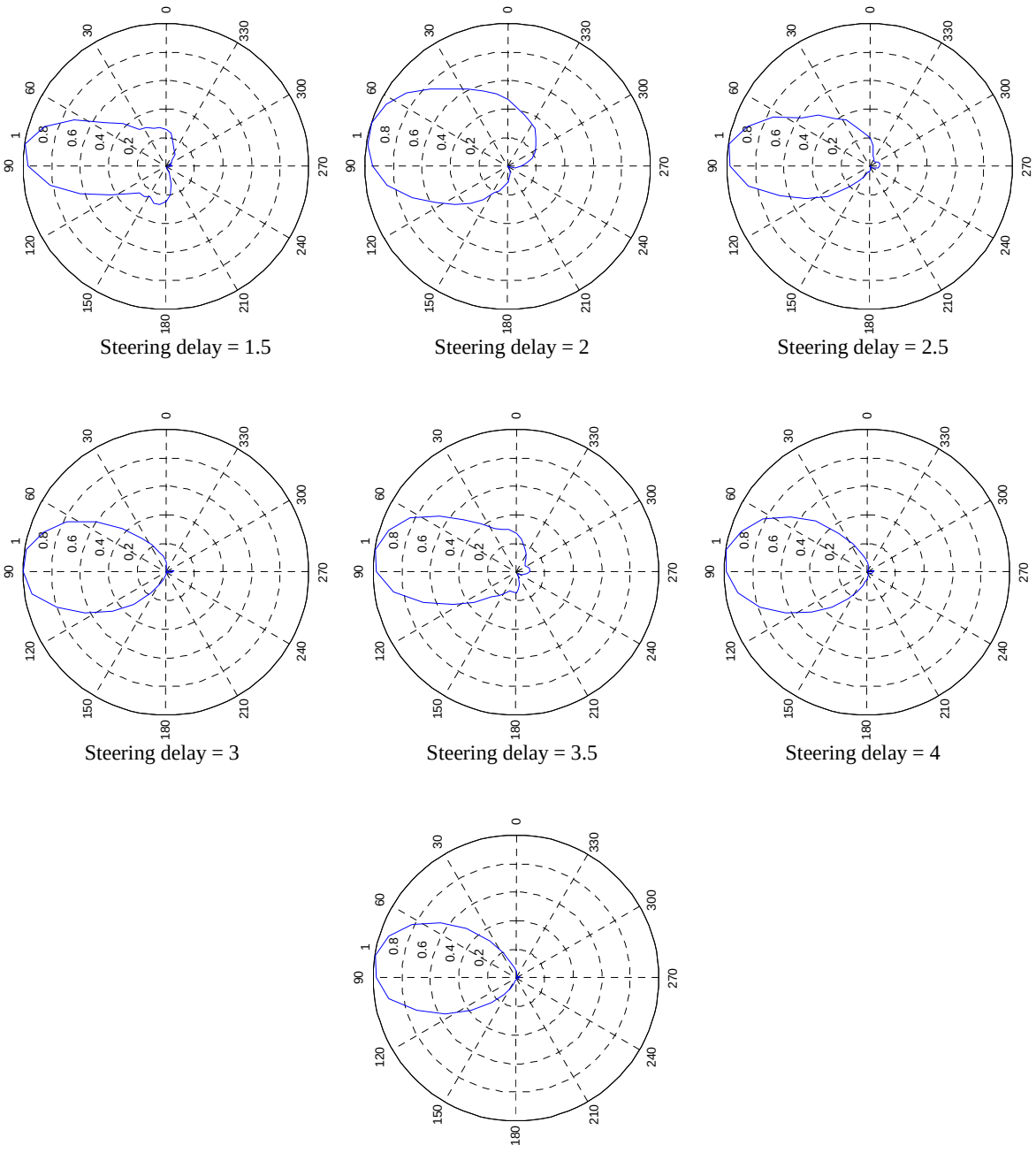
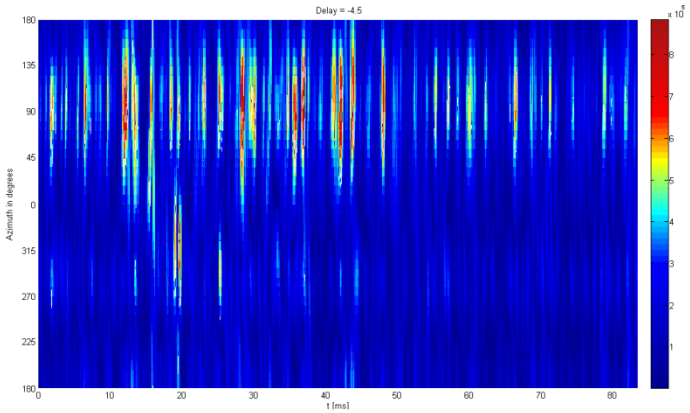
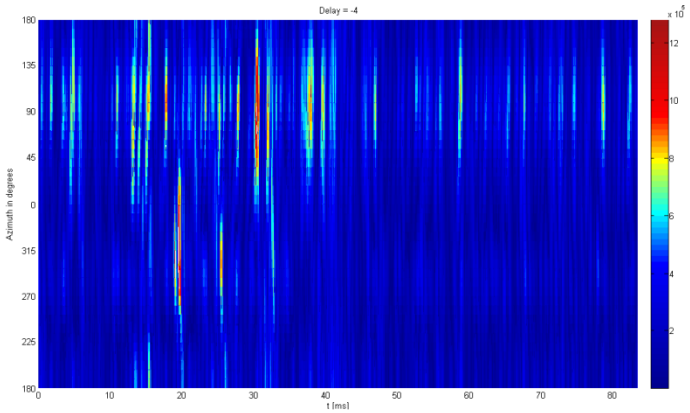



**Figure D.11.** Summary of the received radiation patterns during the sounding phase in a noisy environment using the CSS technique.

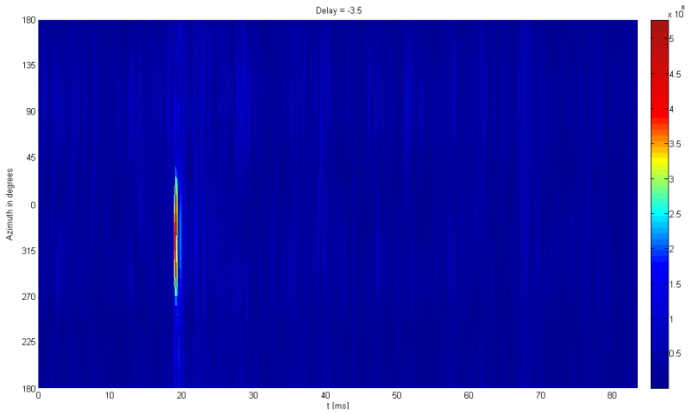
Steering delay = -4.5



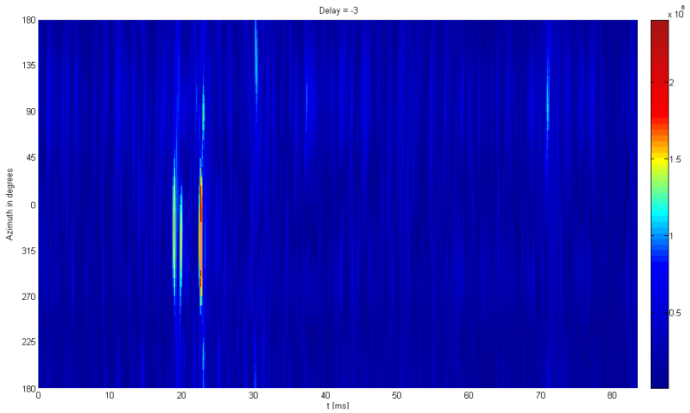

Steering delay = -4



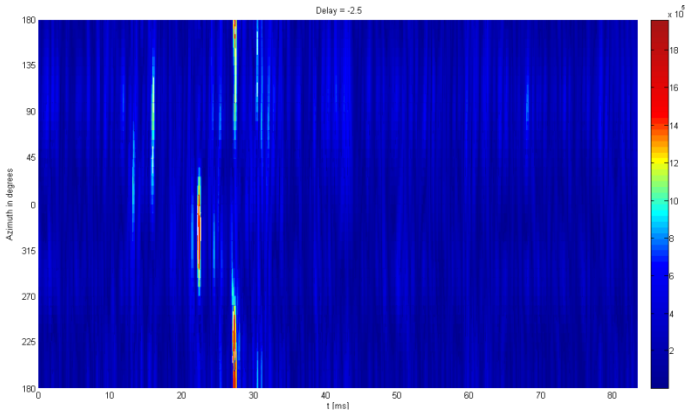

Steering delay = -3.5



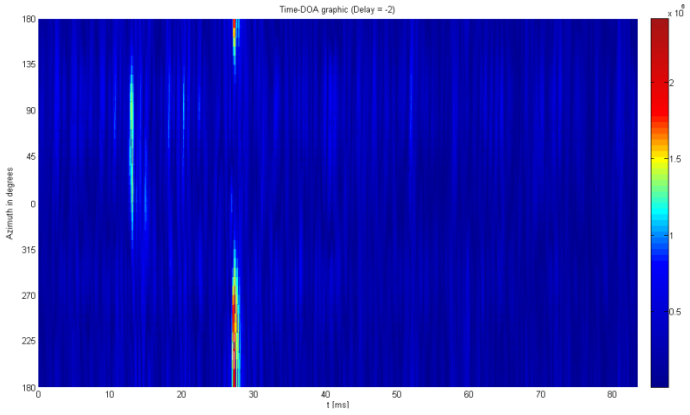

Steering delay = -3



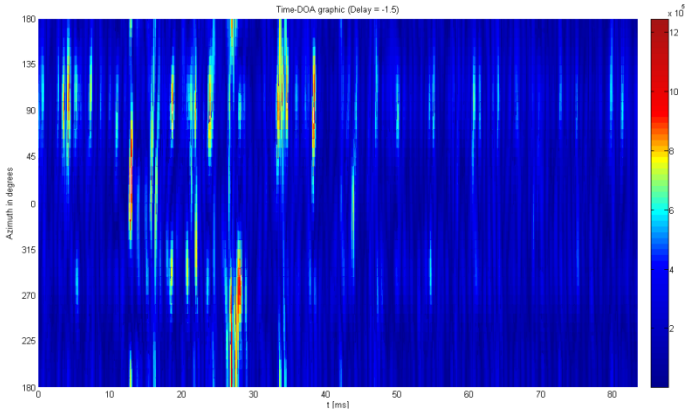
Steering delay = -2.5



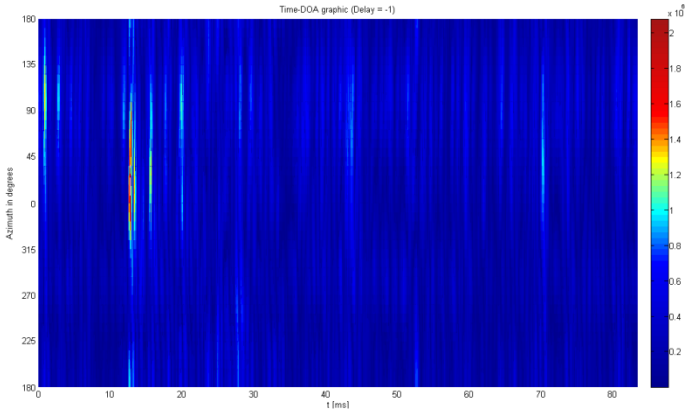

Steering delay = -2



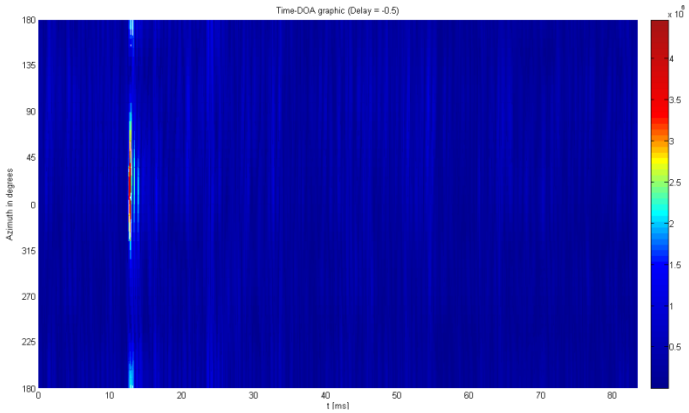

Steering delay = -1.5



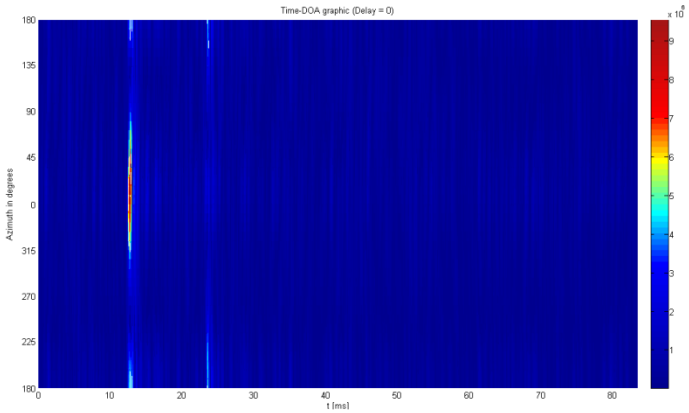

Steering delay = -1



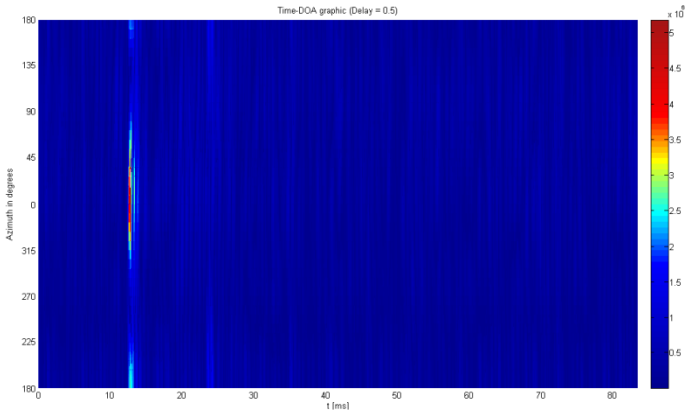
Steering delay = -0.5



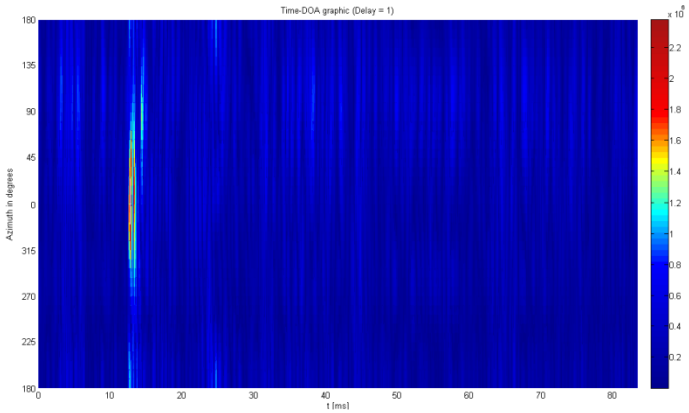

Steering delay = 0



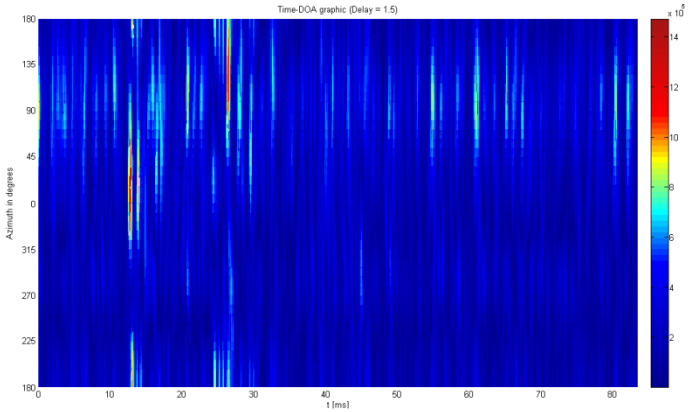

Steering delay = 0.5



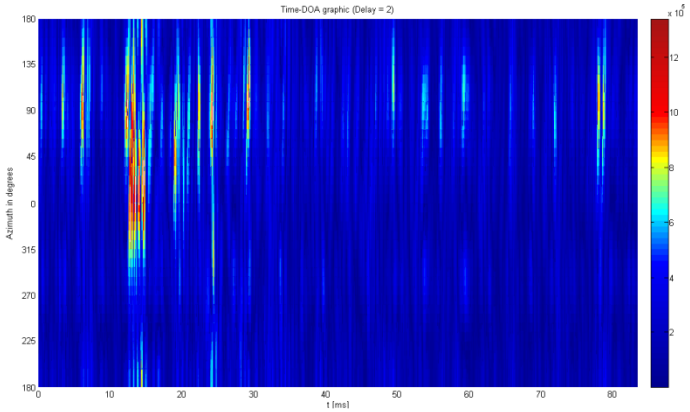

Steering delay = 1



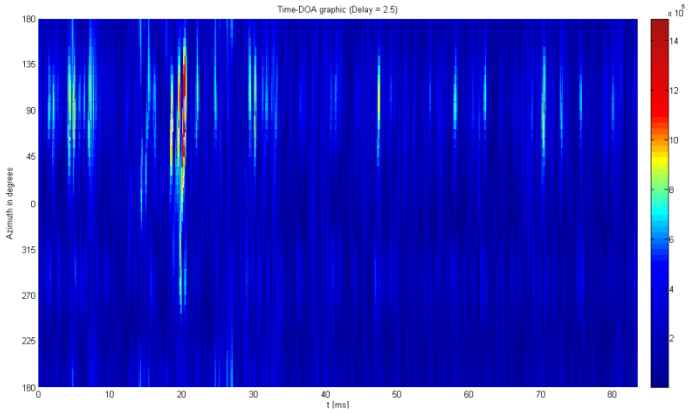
Steering delay = 1.5



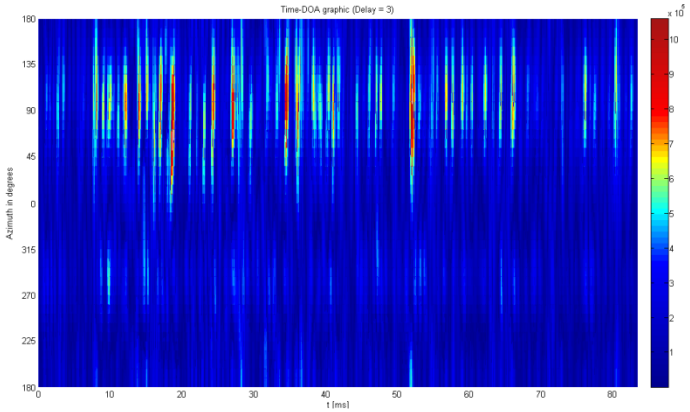

Steering delay = 2



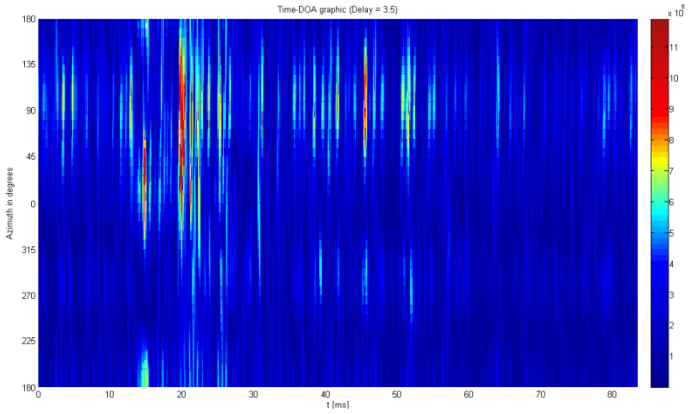

Steering delay = 2.5



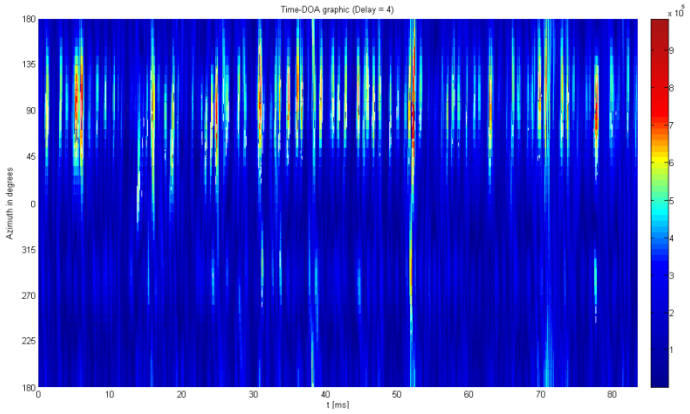

Steering delay = 3



Steering delay = 3.5




Steering delay = 4




Steering delay = 4.5

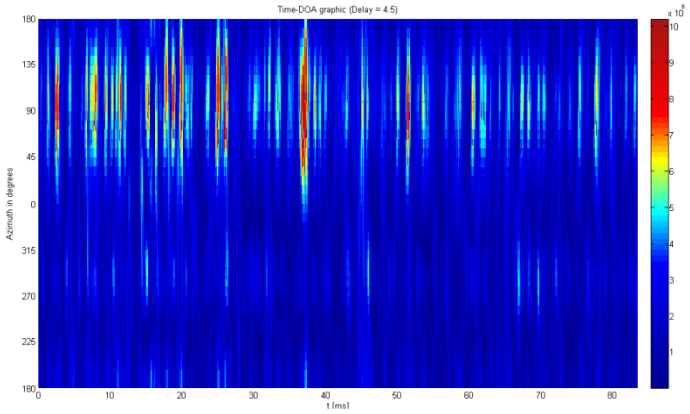



**Figure D.12.** Summary of the propagation delay-DOA plots received during the sounding phase in a noisy environment.